Essays on model uncertainty in financial models

Li, Jing

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Essays on Model Uncertainty in Financial Models

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. E.H.L. Aarts, en Università degli Studi di Torino op gezag van de rector magnificus, prof. dr. G. Ajani, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van Tilburg University op woensdag 17 januari 2018 om 10.00 uur door

JING LI

geboren op 9 december 1984 te Dianbai, China.

PROMOTIECOMMISSIE:

PROMOTORES: Prof. dr. B. Melenberg

Prof. dr. P. Ghirardato

OVERIGE LEDEN: Dr. N. F. F. Schweizer

Dr. P. Cizek

Dr. S. Favaro

Contents

List of Tables

Listing of figures

Acknowledgments

1
Introduction

This dissertation consists of three essays on studying the impact of model uncertainty on financial models. This topic combines the research from both decision theory and control theory. It is motivated by the situation when a practitioner prefers using a specific model (the nominal model) that might not be close to the true data generating process (DGP) even with a sufficiently large sample size. Such a model might be subject to considerable model uncertainty. As a consequence, in such cases it is important to investigate the impact of model uncertainty on the nominal model for decision-making. My research is specifically conducted in the context of asset pricing and financial products.

Model uncertainty, at large, refers to any uncertainty causing a model to deviate from the true DGP. In my research, model uncertainty is classified into two types. The first type is called parameter uncertainty, resulting from lacking sufficient data. It can be eliminated by a sufficiently large sample. The estimated outcome is evaluated by a confidence interval, where the true outcome is expected to fall if the model is correctly specified. However, there could be another type of model uncertainty, called model misspecification uncertainty. When this type of uncertainty appears, only by increasing the data availability might not lead the nominal model to be close to the true DGP. The estimated outcome of the quantity of interest is evaluated by a misspecification interval or an uncertainty interval, which incorporate model misspecification uncertainty, or both types of model uncertainty. In the big data era, lacking sufficient data seems to become a less important issue. Therefore, model misspecification uncertainty becomes more imperative. This dissertation is concerned with this issue and written for better understandings when misspecification uncertainty is present.

Generally speaking, this dissertation focuses on the impact of model uncertainty on bond pricing models. It includes two empirical and one theoretical works. The first two papers are empirical studies. Specifically, the three papers and their connections can be described as follows.

The first paper in Chapter 2, "*An Empirical Investigation of Model Uncertainty, with an Application to Affine Term Structure Models*", introduces firstly the definition and classification of model uncertainty in my research, discusses the evaluation of the impact of model uncertainty, and proposes a misspecification interval to address the range of expectations where the true expectation might fall. Specifically, it investigates, analyzes, and compares the misspecification intervals of the averaged bond yields over time, modelled by several chosen ATSMs different in terms of factor choices. The factors of all ATSMs taken into account are chosen as yield-only ones. The models are estimated by a simple approach (3-step OLS). Applying such a simple setting helps in understanding the fundamental impact of model uncertainty before more complicated considerations are involved. Moreover, in order to understand the transmission of the impact of model uncertainty through bond yields, this paper extends the analysis to the misspecification interval of a max-10-year annuity.

The second paper in Chapter 3, "*Evaluating the Impacts of Bond Pricing Misspecification on Forecasted Funding Ratio*", extends the empirical research in the first paper to study the transmitted impact of bond model uncertainty on funding ratio in a defined benefit pension system. It proposes a prediction interval incorporating model uncertainty to evaluate the range of prediction intervals where the true expectation might fall. There are some differences from the first paper . Firstly, I consider ATSMs with other factors, such as inflation, economic growth and central bank interest rates, etc. Secondly, arbitrage conditions and the vector autoregression assumption are controlled for comparison. Thirdly a few ATSMs' estimation approaches are taken into account to determine the impact of model uncertainty. At last, the postulated economy structure is elaborated with more details of the reality as well. Both empirical research paper show that the impact of model uncertainty could be non-negligible when using a model preferred in practice for the convenience due to its simple structure.

The third paper in Chapter 4, "*The Portfolio Rules and Interest Rates under Model Uncertainty*", contributes to financial theory. It incorporates model uncertainty to study an investment-consumption choice problem, in the spirit of Hansen et al. [41]. In particular, the portfolio rules and assets pricing based on a modified Cox-Ingersoll-Ross investment-consumption model. In this paper, I proposes an approach to address the time consistency issue under commitment, as an alternative to Hansen and Sargent [34]. Under the new approach, the closed-form solutions for the general optimization and equilibrium are obtained. Comparisons are made between some different formulations in literature and under different measures. Calibrations are carried out to explain equity risk premium puzzle based on the established solutions. They show that taking into account misspecification uncertainty is necessary for explaining the puzzle.

In these three pieces of works, I have studied the empirical impact of bond pricing model uncertainty, its transmitted impact on the funding ratio in defined benefit pension scheme, and an investment-consumption choice problem under model uncertainty. These works improve the understandings of the impact of model uncertainty in financial models, as well as raise questions for future research. Related topics can be extended for improvements in policy-making, and econometric and financial theories.

2

An Empirical Investigation of Model Uncertainty, with an Application to Affine Term Structure Models

2.1 Introduction

This paper empirically evaluates the impact of model uncertainty, applied to affine term structure models (ATSMs). Model uncertainty, at large, refers to any uncertainty accounting for a model's deviation from the true data generating process (DGP). To give an example, consider a linear model determined by parameters and explanatory variables under the mean independence assumption, and estimated using a random sample. The most common type of model uncertainty is parameter uncertainty, resulting from sampling error. If the model is correctly specified and sufficient data is available, consistent estimation implies that the estimated model approaches the true DGP. There is also another type of model uncertainty, caused by, for instance, incorrect choice of explanatory variables. This is model misspecification uncertainty. When this type of uncertainty appears, only increasing the sample will not generate an asymptotic outcome equal to the true DGP. Classical econometric theory usually addresses the former, but neglects the latter. Lacking full knowledge of the true DGP, it is sensible for us to take into account model misspecification uncertainty when making decisions based on a nominal model that is preferred to use.

ATSMs are an important class of financial tools to price fixed-income securities. They are used to determine yield curves, and explain yields by a linear relation of factors that are usually assumed to follow a first-order vector autoregression process. There are plenty of suggestions in literature on choosing factors. For instance, Brigo and Mercurio [6] state in their book that traders mostly prefer a single factor, being the short rate. Such a simple model is usually ill-behaved, particularly when used to price bonds with long maturities. As a consequence, the yields derived from such a nominal model might be subjected to a substantial amount of model uncertainty. From this perspective, this paper is motivated to consider the impact of model uncertainty on ATSMs. In this paper, I quantify the impact of model uncertainty by a misspecification interval. This interval is formed by the upper and lower bounds of the expectation derived from an uncertainty set containing models satisfying certain constraints. Depending on situations defining the worst case, a decision maker could make a robust decision based on the upper or lower bound in order to shield from unfavorable extreme situations.

To evaluate the impact of model uncertainty on bond yields, this paper proposes to use a misspecification interval to quantify, in particular, the impact of misspecification uncertainty. A misspecification interval is constructed on an uncertainty set containing all possible models around the nominal model.

This paper constructs the uncertainty set following Schneider and Schweizer [65]'s approach. This approach applies the model confidence set (MSC) [42] procedure to identify a collection of empirically relevant models that perform the best statistically indistinguishably[1]. This provides an approximation where the true DGP is. Measuring the Kullback-Leibler divergence [50] between the nominal model and the MCS, we could decide an appropriate amount of model uncertainty embedded in the uncertainty set, so that the true DGP is most likely to be included. Using the divergence to formulate a constraint, the uncertainty set is the smallest model set constructed to contain the true DGP. The misspecification interval based on this set is considered to be the smallest interval that captures the true expected outcome, when taking into account model uncertainty .

This paper firstly discusses the evaluation theory and explores the empirical application to ATSMs. In the specific context, the quantity of interest is the bond yield with a specific time-to-maturity, averaged over time. The nominal models and the ones in the MCS selection are different only in the factor choices. All factors are individual bond yields with different time-to-maturities. All models are estimated by the 3-step OLS proposed by Adrian et al. [1].

As the uncertainty set is determined, we are able to find out the upper bound and the lower bound models from this uncertainty set, which leads to the best case and worst case bounding the misspecification interval. Glasserman and Xu [27] present a straightforward proposition for such applications. The results are associated with an uncertainty parameter which implies the strength of model uncertainty on the nominal expected outcome. When the models are Gaussian, as in case of ATSMs, the uncertainty parameter can be solved by a simplified closed form.

This paper further discusses the relation between the misspecification intervals and the length of the yield process. I find an increasing relation between them. This is mainly because the multivariate distribution of a longer process will lead to a larger impact of model uncertainty, implying a more conservative attitude when considering a longer period of investment. With a stationarity assumption, one may apply a less conservative strategy by investigating the single-period process.

[1] "Statistically indistinguishable" describes the models that perform equivalently well given the same finite sample, according to some criteria, e.g., a loss function constructed by the log-likelihood function.

I specifically analyze the yields with time-to-maturities 10 months, 30 months, 40 months, 50 months, 60 months, 80 months, 90 months, and 110 months. This allows me to form a yield curve, and analyze the impact of a certain nominal model's model uncertainty on the yield curve as well.

This paper compares the impact of model uncertainty of seven nominal models. These nominal models include two single-factor models, two two-factor models, two three-factor models and a five-factor model with short rates. The results show that the impact of model uncertainty is significant and considerable when using a short rate as a single factor. The bound(s) of the misspecification intervals across maturity are up to 16.11% higher or lower than the nominal yield outcome. The least impacted nominal model is a three-factor model consisting of a short rate, a medium rate, and a long rate. This is in line with the interpretation of the canonical three-factor model [52], which is documented with good empirical fit.

To explore further the impact of model uncertainty of ATSMs, a simple analysis is carried out on a max-10-year annuity that pays 1 Euro once a year during 10 years as long as the annuitant is alive. The annuity uses yields in discounting. The impact of model uncertainty on yields will be transmitted to the annuity value. The results show that, the misspecification interval of the annuity can be as wide as up to 0.44 Euro cents around 9.9422 Euros, when the nominal model uses a short rate as the single factor to estimate the annuity for male at 25; the width of the misspecification interval is similar for female at 25, but around 9.9535. This is much less as compared to the effects on yields, due to the discounting effect.

The primary literature studying model uncertainty is summarized in Hansen and Sargent [36], referring to as robustness for model misspecification. The essential motivation acknowledges that a nominal model is misspecified from the true DGP. A decision maker should confront such misspecification by taking into account alternative possibilities to the nominal one, collected in the uncertainty set, such that he could attain the impact from the uncertainty set, and protect the decisions from downside effects.

There are earlier works in economics and finance incorporating the concern of model misspecification [7, 39, 40]. These studies focus on the classical models of consumption, savings, economic growth, and market price of risk, by regarding model misspecification as "Knightian uncertainty" [49], and applying the max-min philosophy with multiple priors in the spirit of Gilboa and Schmeidler [26]. A strand of literature developed by Anderson et al. [2], Hansen and Sargent [32, 33, 37], and Hansen and Sargent [38] eventually establishes their framework to evaluate this kind of uncertainty. Remarkable applications to financial models include Maenhout [57] who studies portfolio allocation, Maenhout [58] who captures the uncertainty effect on bond risk premiums, and Gagliardini et al. [24] who develop interest rates model with uncertainty.

However, in Hansen and Sargent [36]'s framework, the misspecification refers to the case that the nominal model (the approximating model, in their text) does not take the right forms (resulting in a distorting model, in their text). However these models are statistically indistinguishable from the true DGP. The uncertainty set is constructed on this assumption [2]. This implies two things. One is that the uncertainty considered in their framework is actually the parametric uncertainty in my paper's classification, curable by improving data availability. The misspecification uncertainty is not taken into account. The other implication is that, applying their misspecification definition, the nominal model could possibly be a complicated model difficult to implement, suppressing the preference in practical use.

Schneider and Schweizer [65] relax the concerns aforementioned. The uncertainty set in this framework is constructed without restricting whether the models are indistinguishable. In other words, it allows model misspecification uncertainty; the nominal model can be statistically distinguishable from the true DGP. This framework is more appropriate to consider the impact of model uncertainty of simple models. This framework is closely related to the literature that sets up the uncertainty set in terms of divergence between distribution measures [4, 5, 22].

Related literature in finance about the factors used in ATSMs is enormous. The classical ones are one-factor models by Vasicek [69] and Cox et al. [12]. They are applied broadly for the simplicity in practice, although they may not provide good empirical fits. There are also well-known three-factor models [15, 52]. The factors proposed to use include principle components [9, 45], latent yield factors and (unspanned) macroeconomic factors [14, 16, 19, 46, 54], etc. For illustrative purpose, this paper will not take into account complicated factors, but only yields with various maturities.

In literature, applying Schneider and Schweizer [65]'s framework to study ATSMs is limited. This paper contributes to develop the empirical method to evaluate the impact of model uncertainty for ATSMs, and provide empirical analysis to understand the fundamental impact of model uncertainty in this context.

The remainder of the paper is organized as follows. Section 2.2 provides the theoretical explanation and discussions of the empirical evaluation method. Section 2.3 presents the main empirical analysis and findings. The last section concludes.

2.2 Some Model Uncertainty Evaluation Theory

In statistics, a model is typically defined as a probability distribution p of $y \in \mathbb{R}^{s_1 \times 1}$, conditional on variables $X \in \mathbb{R}^{s_2 \times 1}$, where p is characterized by the unknown parameter vector φ, to be estimated by a finite sample. Let the true DGP be given by p_0. In practice, one selects a specific model, denoted by p_{NOM}, to work with.

How well p_{NOM} is able to capture p_0 can be described by a so-called confidence set $\mathcal{P}_C$. Intuitively, the confidence set $\mathcal{P}_C$ is a set of probability distributions around p_0 which we do not want to shrink further, because otherwise the probability of leaving p_0 outside $\mathcal{P}_C$ is considered to be too high. Typically, $\mathcal{P}_C$ is not exactly equal to $\{p_0\}$, but the models within are considered to be statistically indistinguishable from p_0. Should sufficient data be provided and the model set-up be correct, the confidence set would shrink to $\{p_0\}$.

There are many ways to construct $\mathcal{P}_C$. Depending on the objects to tackle, the following two are closely relevant to the subject addressed in this paper. The first one is the standard approach based on the ordinary confidence set. Consider a set of models $\mathcal{P}_\Phi$ such that for some $\varphi_0 \in \Phi$ we would have $p_{\varphi_0} = p_0$, and a consistent estimator $\widehat{\varphi}$ such that $\widehat{\varphi} \xrightarrow{p} \varphi_0$ with an asymptotic distribution $\sqrt{n}\left(\widehat{\varphi} - \varphi_0\right) \xrightarrow{d}$

$\mathcal{N}(0, V_{\varphi_0})$ as the sample size $n \to \infty$. Then the confidence set is constructed around $p_{\widehat{\varphi}}$, given by $\mathcal{P}_C = \{ p_\varphi \in \mathcal{P}_\Phi | \, \varphi \in \Phi_C \}$ where Φ_C is a confidence set of some level around $\widehat{\varphi}$.

The second approach is based on the idea of the model confidence set (MCS) by Hansen et al. [42]. Generally, it starts from some set $\{p_a | \, a \in A\}$, such that (preferably) there exists an $a_0 \in A$ for which $p_{a_0} = p_0$. Then based on data and using a sequential procedure, the models in A yielding empirical outcomes sufficiently worse than the others are excluded, until achieving a subset $A' \subset A$ so that no further exclusion can be motivated. The confidence set is then $\mathcal{P}_C = \{p_a | \, a \in A'\}$.

This paper distinguishes two cases concerning the relationship between p_{NOM} and $\mathcal{P}_C$. The first one is $p_{NOM} \in \mathcal{P}_C$. If so, p_{NOM} is a well-performing model in terms of data-fitting for p_0. A common instance is $p_{NOM} = p_{\widehat{\varphi}}$. The only reason why p_{NOM} diverges from p_0 is due to insufficient data. This uncertainty is referred to as parameter uncertainty. The second one is $p_{NOM} \notin \mathcal{P}_C$. A typical example is $p_{NOM} \notin \mathcal{P}_\Phi$ or $p_{NOM} \notin \{p_a | \, a \in A\}$. In this case, the model uncertainty of p_{NOM} does not merely come from a lack of data, i.e., parameter uncertainty. It includes also another type of uncertainty that cannot be corrected only be increasing data availability. This is called model misspecification uncertainty. In this case, this type of uncertainty is imperative because of the non-negligible statistical distinction from p_0. The model uncertainty I consider is composed of these two types.

When $p_{NOM} \in \mathcal{P}_C$, it itself is deemed as an adequate estimator for p_0 based on the available data. The uncertainty set, denoted by $\mathcal{P}_U$, can be established like $\mathcal{P}_C$, in particular, if $p_{\widehat{\varphi}} = p_{NOM}$. This is in fact the uncertainty set by Anderson et al. [2], addressed to Hansen and Sargent [36]'s framework. Within this set, the models are statistically indistinguishable, though might be concretely different. These models are considered equivalent in describing p_0.

In financial practice, the assumption that $p_{NOM} \notin \mathcal{P}_C$ is more realistic. Practically, it is very common to choose a simple p_{NOM} to reduce the computational burden, even if it does not perform well enough compared to other models. Moreover, the real financial environment might be dynamic, and change the $\mathcal{P}_C$. As a consequence, model misspecification uncertainty could become more significant.

Taking into account model uncertainty, it is important to construct an uncertainty set $\mathcal{P}_U$ for p_{NOM}. $\mathcal{P}_U$ collects all possible models under some conditions. For instance, it collects all the models for which the model uncertainty with respect to p_{NOM} is not greater than a certain amount. The impact of model uncertainty can be quantified by the outcomes ruled by the models within $\mathcal{P}_U$.

The remaining part of this section will first discuss the ATSMs estimation. It is then followed by the tailored MCS procedure to obtain the model confidence set as $\mathcal{P}_C$. The last part works on the formation of the uncertainty set $\mathcal{P}_U$, and discusses the evaluation of the impact of model uncertainty generally and specifically for ATSMs.

2.2.1 The ATSMs Estimation

The classical ATSMs consider a bond yield $y_t^{(\tau)}$ with time-to-maturity τ, that is affine in a vector of factors X_t following a VAR(1) process, given by

$$y_t^{(\tau)} = -\frac{1}{\tau}\left(A_\tau + B_\tau' X_t + u_t^{(\tau)}\right), \tag{2.1}$$

$$X_{t+1} = \mu + \Psi X_t + v_{t+1}, \tag{2.2}$$

where A_τ, B_τ, μ and Ψ are parameters. The scalar $u_t^{(\tau)}$ and the vector v_{t+1} are normally distributed with mean 0, and variance Σ_u and Σ_v, respectively. The variances of X_t and $y_t^{(\tau)}$ are denoted by Σ_X and Σ_y, respectively. In most research, the arbitrage-free condition is also imposed, and the parameters of A_τ and B_τ are recursively determined by a set of parameters including μ and Ψ.

The conventional estimation method is the maximum likelihood estimation, which is difficult to implement. Adrian et al. [1] propose an approach based on linear regressions: the 3-step OLS approach. In this approach, no technique more complicated than ordinary least squares (OLS) is needed. See the original paper for the technical estimation details. They use this method to investigate bond risk premia by four- and five-factor specifications of principal components. They find that this method outperforms Cochrane and Piazzesi [10]'s model in an out-of-sample exercise.

The VAR(1) process of the factors X_t in (2.2) implies zero measurement error. This is a strong assumption, if the factors are individual bond yields instead of principal components. In a data set, there typically exist correlations between individual bond yields, possibly causing considerable measurement errors. Principal components, on the contrary, would mitigate such worries. However, the application of principal components requires a collection of vast bond yields, and additional estimation. Moreover, the simple choices in practice are the main reason to motivate this study of the impact of model uncertainty. Therefore, this paper is inclined to maintain the simplest construction of models, and use only individual bond yields rather than principal components as factors.

2.2.2 The Model Confidence Set

To specify the confidence set $\mathcal{P}_C$ of the true DGP p_0 from a large collection of ATSMs, one approach is the model confidence set procedure [42]. This procedure chooses a set of indistinguishable best-performing model(s) $\mathcal{M}^*$ from a collection of potentially empirically relevant models $\mathcal{M}^\circ$, at a specific confidence level α. It is analogous to a conventional confidence interval. Instead of evaluating estimates, MCS targets on a class of models. This selection procedure is limited to the information provided by the data. Sufficiently informative data will result in only one best model left, while less informative data will correspond to more remaining models as it is not powerful enough to distinguish them.

Specifically, the MCS procedure goes as follows. Under a chosen significance level α, an equivalence test is implemented based on the values of a loss function associated with the models in $\mathcal{M}_r \subset \mathcal{M}^\circ$, where $\mathcal{M}_r$ is the model set in the selection round r before the whole procedure ends. The loss function qualifies a model's performance: the greater its value is, the worse the model explains the data. The null hypothesis assumes that the expectations of all loss functions are equal. In other words, all models are statistically indistinguishable at the significance level α. If the p-value associated with the test is smaller than α, the null hypothesis is rejected. Then the model with the largest loss function value, $e_{\mathcal{M}_r}$, will be removed. The set with the remaining models will tested for a new round. This procedure continues until the null is no longer rejected.

The ATSMs given by (2.1) are linear models. Denote by $y^{(\tau)}$ a vector containing a T-period univariate time series of the bond yield with time-to-maturity τ, i.e., $y^{(\tau)} = \left(y_1^{(\tau)}, \cdots y_T^{(\tau)}\right)'$. $y^{(\tau)}$ is conditional on a set of factors collected in the vector X. The specific factors used in model j with $j = 1, \cdots, m_0$ is chosen by $S_j X$, where S_j is a model-specific selection matrix of 0s and 1s.[2] Consider an initial model set $\mathcal{M}^0$ consisting of a finite number of potentially empirically relevant models. Represent these models by the probability distribution p_{Z_j, φ_j}, where $Z_j = \{S_j X, y^{(\tau)}\}$ collects the observations of the factors X and the bond yield $y^{(\tau)}$, and $\varphi_j = \{A_{\tau,j}, B_{\tau,j}, \Sigma_{u,j}\}$ collects the parameters characterizing model j.

The loss function is constructed by the conditional log-likelihood function $\log l\left(\varphi_j\right)$. In essence, $\log l\left(\varphi_j\right) = \log p_{\left(y^{(\tau)} | S_j X, \varphi_j\right)}$, applying the log of the conditional distribution. The ATSMs considered in this paper are a family of one-dimensional linear regression models with normal distribution. For simplicity, we use instead the quasi-log-likelihood function, given by

$$\log l\left(\varphi_j\right) = -\frac{T}{2} \log \Sigma_{u,j} - \frac{1}{2} \sum_{t=1}^{T} \frac{\left[y_t^{(\tau)} + \left(A_{\tau,j} + B'_{\tau,j} S_j X_t\right) / \tau\right]^2}{\Sigma_{u,j}}, \tag{2.3}$$

The loss function is thus defined by

$$Q_j\left(Z_j, \varphi_j\right) = -2 \log l\left(\varphi_j\right) = T \log \Sigma_{u,j} + \sum_{t=1}^{T} \frac{\left[y_t^{(\tau)} + \left(A_{\tau,j} + B'_{\tau,j} S_j X_t\right) / \tau\right]^2}{\Sigma_{u,j}}. \tag{2.4}$$

[2] Suppose $X = [x_1, x_2, x_3, x_4]'$ for instance. For a model indexed j using only the factors x_1 and x_3, the selected factors can be expressed by $S_j X$ where $S_j = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$.

The elements in $\mathcal{M}_r$ are ranked according to $Q_j\left(Z_j, \varphi_j\right)$ in descending order, giving $\mathcal{M}_r = \{e_{\mathcal{M}_r}, \cdots, e_{\mathcal{M}_{m_0}}\}$ where $r = 1, 2, \cdots, m_0 - 1$. The test procedure starts from $\mathcal{M}_1 = \mathcal{M}^0$.

The test in the MCS procedure needs resampling to construct the statistic. Since both the bond yield to model $y^{(\tau)}$ and the factors X are assumed to be strictly stationary stochastic processes, in order to capture the serial relation of these variables, the test is implemented by the moving-block bootstrap approach, following Hansen et al. [42]'s argument. This approach retains the time serial correlation because it draws resamples by a length of L successive periods from the original sample. For testing implementation details, see the original paper.

The p-value of the set $\mathcal{M}_r$ in the rth-round is denoted by p_r-val. When p_r-val $>$ α, the null hypothesis of model indifference is not rejected. The procedure stops, and $\mathcal{M}_r$ is taken as the MCS $\mathcal{M}^*_{1-\alpha}$, or $\mathcal{P}_C$, at the confidence level $1 - \alpha$.

$\mathcal{M}^*_{1-\alpha}$ contains the indistinguishable model(s) superior to the eliminated ones. With sufficiently powerful data, the more the potentially empirically relevant models are included in the initial set $\mathcal{M}^0$, the more reliable the MCS describing p_0 is. The parameter α also affects the selection. The higher it is, the stricter the threshold is, and the less likely the worst-performing model can survive. It also implies that the practitioners are less tolerant towards parameter uncertainty.

2.2.3 The Uncertainty Set and Its Impact Evaluation

The choice of a simple nominal model for practical reasons may lead to a situation where the nominal model is not in the MCS, i.e., $p_{NOM} \notin \mathcal{P}_C$. The impact of model uncertainty is evaluated by a misspecification interval, obtained from the $\mathcal{P}_U$ containing a certain amount of model uncertainty. In the following subsection, the construction of $\mathcal{P}_U$ will firstly be explained. Then I will show how $\mathcal{P}_U$ is linked to an uncertainty parameter, and how the uncertainty parameter determines the misspecification interval. In the end, I will apply the theory to analyze ATSMs.

The Construction of $\mathcal{P}_U$

In Schneider and Schweizer [65]'s proposal, ideally $\mathcal{P}_U$ should contain all the best-performing models in order to justify its inclusion of p_0, and thus, an appropriate amount of model uncertainty to consider. Hence, it requires that (1) $\mathcal{P}_U \supset \mathcal{P}_C$ (now the MCS) so that the uncertainty amount is constraint to be large enough to account for all models in $\mathcal{P}_C$. (2) Moreover, $\mathcal{P}_U$ should contain any model satisfying the the uncertainty amount constraint between p_{NOM} and $\mathcal{P}_C$. In other words, we do not need to consider the models with uncertainty amount larger than the constraint.

A typical indicator to quantify model uncertainty is a divergence D, which measures the discrepancy between two models. Applying D, $\mathcal{P}_U$ is constructed by

$$\mathcal{P}_U = \left\{ p_a \middle| D\left(p_a \middle|\middle| p_{NOM}\right) \leq \kappa^*, \kappa^* = \sup\left[D\left(p_{a'} \middle|\middle| p_{NOM}\right) \middle| p_{a'} \in \mathcal{P}_C \right] \right\}, \tag{2.5}$$

where $D\left(p_{a'}\middle|\middle|p_{NOM}\right)$ measures the amount of model uncertainty of p_{NOM} from any $p_{a'}$ in $\mathcal{P}_C$, and κ^* is the largest amount given by the models in$\mathcal{P}_C$. Figure 2.2.1 illustrates the construction. One may consider p_{NOM} an element in the model space. In this model space, there are other probability models, e.g., marks like $\times$. Given available data, we could identify a $\mathcal{P}_C$ containing models performing the best indistinguishably in the sense of statistics. It is *MCS* if applying the MCS selection procedure. The set $\mathcal{P}_C$ is considered to provide an approximation where the true DGP is. The amount of model uncertainty ought to be taken into account is determined by the largest divergence between p_{NOM} and $\mathcal{P}_C$, denoted by κ^*. In the end, one may think of $\mathcal{P}_U$ as a sphere centering around p_{NOM} with the radius κ^*, and $\mathcal{P}_C$ internally tangent to $\mathcal{P}_U$. The internal tangency corresponds to the first condition, while the sphere with the radius κ^* corresponds to the second condition. The part within $\mathcal{P}_C$ reflects parameter uncertainty, while the part between $\mathcal{P}_C$ and p_{NOM} corresponds to model misspecification uncertainty. The uncertainty set $\mathcal{P}_U$ becomes the constraint within which we evaluate the impact of model uncertainty if using p_{NOM}.

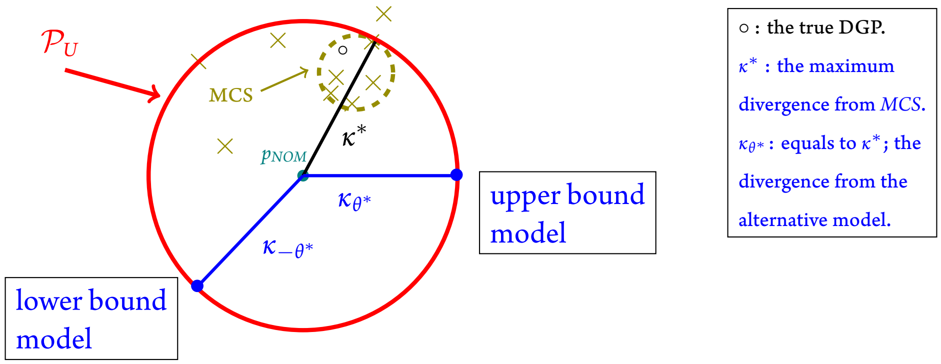


Figure 2.2.1: Diagram of the model uncertainty set

This paper uses one of the most common divergences, namely, the Kullback-Leibler (KL) divergence. I continue to apply the selection matrix S_j to X for a model j,[3]. Moreover, I simplify the notation $y^{(\tau)}$ to be y in Section 2.2.3, because the maturity τ is not relevant here. Generally, the nominal model conditional on $S_{NOM}X$ is $p_{NOM}(y|S_{NOM}X)$, and an alternative model j conditional on S_jX is $p_j(y|S_jX)$. Assume that the deselected factors do not play a role, i.e. $p_{NOM}(y|X) = p_{NOM}(y|S_{NOM}X)$ and $p_j(y|X) = p_j(y|S_jX)$. Considering y is a vector, these models are joint distributions of $y_t^{(\tau)}$. Moreover, these models are assumed to be absolutely continuous with respect to the Lebesgue measure. The KL divergence of

[3]When j refers to a nominal model, the subscript is NOM; when j refers to the model with the maximum divergence in the MCS, it is marked by $*$. Unless specially specified, j in the sequel indicates an alternative model in $\mathcal{P}_U$

$p_{NOM}(y|X)$ from $p_j(y|X)$ is defined explicitly by [4] [5]

$$\begin{aligned} & D\left[p_j(y|X)\middle|\middle|p_{NOM}(y|X)\right] \\ :=& \int\int \frac{p_j(y|X)}{p_{NOM}(y|X)} \log\left[\frac{p_j(y|X)}{p_{NOM}(y|X)}\right] p_{NOM}(y|X)p\left(X\right) dxdy \\ =& \mathbb{E}_{NOM}\left(\widetilde{m}_j \log \widetilde{m}_j\right), \end{aligned} \tag{2.6}$$

where $\widetilde{m}_j := p_j\left(y|X\right)/p_{NOM}\left(y|X\right)$ is a well-defined likelihood ratio, commonly called Radon-Nikodym derivative (hereinafter, RN derivative).

The KL divergence in (2.6) implies the amount of information loss if $p_{NOM}(y|X)$ is used to approximate $p_j(y|X)$, or the information gain if $p_j(y|X)$ is used instead of $p_{NOM}(y|X)$. In some contexts, the KL divergence is also called the relative entropy of $p_j(y|X)$ with respect to $p_{NOM}(y|X)$.

By the definition of $\widetilde{m}$, an alternative conditional probability distribution can be easily obtained from a nominal one by a change of measure $p_j(y|X) = \widetilde{m}_j \cdot p_{NOM}(y|X)$. The definition of (2.6) is based on this relation. Sometimes, this change of measure is quite convenient in practice. For instance, the divergence D in (2.6) with respect to $p_{NOM}\left(y|X\right)$ can be obtained with respect to $p_j\left(y|X\right)$; see the footnote 5.

Typically, we are interested in the quantity of an expectation $\mathbb{E}_j\left[g(y)\right]$ for some function g of y (e.g., the average of bond yield over time). It depends on the marginal distribution, $p(y)$, instead of the conditional distribution $p(y|X)$.[6] For this reason, we look for the RN derivative with respect to $p(y)$. Specifically in this paper, $g : \mathbb{R}^T \to \mathbb{R}$, where T is the number of periods in y process. The aforementioned

[4] $\mathbb{E}_{NOM}[\cdot]$ is denoted as the expectation taken under the nominal probability measure.

[5] Alternatively, it can be defined as

$$D\left[p_j(y|X)\middle|\middle|p_{NOM}(y|X)\right] := \int\int \log\left[\frac{p_j(y|X)}{p_{NOM}(y|X)}\right] p_j(y|X)p\left(X\right) dxdy = \mathbb{E}_j(\log \widetilde{m}_j),$$

where $\mathbb{E}_j[\cdot]$ denotes the expectation ruled by an alternative probability measure $p_j(y|X)p\left(X\right)$, or more fundamentally, the conditional measure $p_j(y|X)$.

[6] $\mathbb{E}_j\left[g(y)\right] = \mathbb{E}_j\left\{\mathbb{E}_j[g(y)|X]\right\}$, where $\mathbb{E}_j\left[g\left(y|X\right)\right]$ is ruled by $p_j(y|X)$.

change of measure is modified for using $p_j(y)$[7], i.e., $p_j(y) = m_j \cdot p_{NOM}(y)$, and thus the RN derivative is given by $m_j := p_j(y)/p_{NOM}(y)$. The corresponding KL divergence $D\left[p_j(y)\middle|\middle|p_{NOM}(y)\right]$ between $p_j(y)$ and $p_{NOM}(y)$ is thus equal to

$$\begin{aligned} D\left[p_j(y)\middle|\middle|p_{NOM}(y)\right] &= \int \frac{p_j(y)}{p_{NOM}(y)} \log\left[\frac{p_j(y)}{p_{NOM}(y)}\right] p_{NOM}(y)\, dy \\ &= \mathbb{E}_{NOM}\left(m_j \log m_j\right). \end{aligned} \tag{2.7}$$

In the sequel, when calculating the KL divergence I can apply the definition in (2.7) in terms of marginal distributions.

Denote by $\kappa > 0$ some given number. $\mathcal{P}_U$ is composed of all models from which the divergence is not larger than κ. Or, equivalently, it can be considered as a set $\mathcal{P}_\kappa$ of likelihood ratios m for which $\mathbb{E}\left(m \log m\right) \leq \kappa$. To satisfy the condition $MCS \subset \mathcal{P}_U$, the largest divergence from the models in MCS is applied, denoted by κ^*. The MCS model corresponding to κ^* is denoted by $p_*(y)$, whose factor choice is S_*X. Then, the proper equivalence of $\mathcal{P}_U$ is given be

$$\mathcal{P}_{\kappa^*} = \left\{m_j : \mathbb{E}_{NOM}\left(m_j \log m_j\right) \leq \kappa^*, \kappa^* = D\left[p_*(y)\middle|\middle|p_{NOM}(y)\right]\right\}, \tag{2.8}$$

Taking into account the time series of bond yields $y^{(\tau)}$, and the pertaining models are multivariate Gaussian distributed with T-dimensioned, where T is the number of periods in consideration. The KL divergence κ between two models is given by

$$\kappa = \frac{1}{2}\left[\log\frac{|\Sigma_2|}{|\Sigma_1|} - T + tr(\Sigma_z^{-1}\Sigma_1) + (\mu_2 - \mu_1)' \Sigma_2^{-1}(\mu_2 - \mu_1)\right], \tag{2.9}$$

where $\mu_1 \in \mathbb{R}^{T\times 1}$ and $\mu_2 \in \mathbb{R}^{T\times 1}$ are the means of $y^{(\tau)}$ under the probabilities

[7]By multiplying the denominator and the nominator of $\widetilde{m}$ by $p(X)$, one can observe that the definition in (2.6) is also applicable to the unconditional divergence $D\left[p_j(y,X)\middle|\middle|p_{NOM}(y,X)\right]$ between $p_j(y,X)$ and $p_{NOM}(y,X)$. So is it for the change of measure. Then, integrating with respect to X the change of measure between $p_j(y,X)$ and $p_{NOM}(y,X)$ provides the change of measure between the marginal distributions $p_j(y)$ and $p_{NOM}(y)$.

of $p_j(y)$ and $p_{NOM}(y)$ respectively, and Σ_1 and Σ_2 are the corresponding variances. Typically, κ increases in the dimension T. The amount of model uncertainty κ^* is chosen to be the largest one from all κs between the nominal model and the MCS models.

The Evaluation of the Impact of Model Uncertainty

One significant reason for using the KL divergence is because of its definition by the RN derivative, while the RN derivative is the key to change a measure by $p_j(y) = m_j \cdot p_{NOM}(y)$. This helps in evaluating $\mathbb{E}_j[g(y)]$ under $p_{NOM}(y)$, i.e.,

$$\mathbb{E}_j[g(y)] = \mathbb{E}_{NOM}\left[m_j \cdot g(y)\right]. \tag{2.10}$$

Under certain mathematical conditions stated in Glasserman and Xu [27], the quantity $\mathbb{E}_j[g(y)]$ is bounded by construction. In other words, from $\mathcal{P}_U$ of κ^*, there exist an upper bound and a lower bound of $\mathbb{E}_j[g(y)]$. These bounds of the expectation form the misspecification interval of $\mathbb{E}_{NOM}[g(y)]$.

Suppose c is an indicator that equals 1 for the upper bound, and -1 for the lower bound. Following the proof in Glasserman and Xu [27], the bound models subject to the set $\mathcal{P}_{\kappa^*}$ in (2.8) are obtained by bounds of $\mathbb{E}_{NOM}[m \cdot g(y)]$,

$$\sup_{m \in \mathcal{P}_\kappa} \mathbb{E}_{NOM}[cm \cdot g(y)]. \tag{2.11}$$

Proposition 3.1 (i) in Schneider and Schweizer [65] provides the solution for the upper bound. The following proposition complements this proposition by including also the lower bound solution.

Proposition 1. *For a model of y given by $p_{NOM}(y)$, suppose there exists a $\theta > 0$ such that the expectation $M(c\theta) = \mathbb{E}_{NOM}\{\exp[c\theta \cdot g(y)]\} < \infty$, and that $M(c\theta) \to \infty$ for $\theta \to \overline{\theta}$ where $c = \pm 1$, and $\overline{\theta} = \sup\{\theta : M(c\theta) \to \infty\}$. Then there exists a $\theta > 0$ such that the probability measure $p_{\theta,NOM}(y)$, given by*

$$p_{\theta,NOM}(y) = \frac{\exp[c\theta \cdot g(y)]}{\mathbb{E}_{NOM}\{\exp[c\theta \cdot g(y)]\}} p_{NOM}(y), \tag{2.12}$$

attains $\mathbb{E}_{c\theta}\left[g(y)\right] = \mathbb{E}_{NOM}\left[p_{\theta,NOM}(y)/p_{NOM}(y)\cdot g(y)\right]$ *which is the upper bound for* $c = 1$*, and the lower bound for* $c = -1$*.* $p_{\theta,NOM}(y)$ *is a density functions of* y *under a* θ*-associated model changed from the nominal model, respectively.*

Proof. Optimize (2.11) by solving the Lagrangian dual form,

$$\inf_{\theta>0}\sup_{m}\mathbb{E}_{NOM}\left[cm\cdot g(y) - \frac{1}{\theta}\left(m\log m - \kappa^*\right)\right], \tag{2.13}$$

where θ is in fact a reciprocal of an Lagrangian multiplier that implies the degree of influence of the restriction $\mathbb{E}\left(m\log m\right) \leq \kappa^*$. The inner $\sup_m$ problem is the Lagrangian dual function, solved by the optimal

$$m^*_\theta = \frac{\exp\left[c\theta\cdot g(y)\right]}{\mathbb{E}_{NOM}\left\{\exp\left[c\theta\cdot g(y)\right]\right\}}, \qquad \text{for some } \theta > 0, \tag{2.14}$$

provided that the expectation in the denominator is finite. This actually has the form $m^*_\theta \propto \exp\left[c\theta\cdot g(y)\right]$. The bound probability is given by the transformation with m^*_θ, i.e., $p_{\theta,NOM}(y) = m^*_\theta\cdot p_{NOM}\left(y\right)$. It turns out that the bound expectations $\mathbb{E}_{NOM}\left[m^*_\theta\cdot g(y)\right]$ following (2.10) are associated with θ, denoted by $\mathbb{E}_{c\theta}\left[g(y)\right]$.

□

The outer $\inf_{\theta>0}$ problem in (2.13) looks for the θ that optimizes the bound expectations under the bound models. To solve for the optimal θ, one can substitute (2.14) back into (2.7). The optimum θ^* is obtained by calibrating θ according to

$$\kappa_\theta = \mathbb{E}\left(m^*_\theta\log m^*_\theta\right), \tag{2.15}$$

such that $\kappa_\theta = \kappa^*$. Recall the diagram in Figure 2.2.1. It shows that the bound models are pertaining to the parameter θ which results in $\kappa_\theta = \kappa^*$.

Because θ^* implies the strength of the model uncertainty constraint, it is called the uncertainty parameter. With the solution θ^*, both $m^*_{\theta^*}$ and $\mathbb{E}_{c\theta^*}\left[g(y)\right]$ can be recovered for the amount of model uncertainty κ^*. In other words, the misspecification interval, the impact of model uncertainty, is retrieved.

The function $g(y)$ is a function of choice by the investigator. Typically, for the yields following an arbitrary distribution , one may choose to investigate the first or second moment of the yields, or even higher moments. In this way, one could capture more features of a distribution under model uncertainty. As we assume normal distribution, what characterize the distribution are the first and second moment. This paper specifically looks into the first moment, i.e. $g(y) = \bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$, the yields averaged over time. By this function, one may think of $\bar{y}$ as the bond yield for a specific period for $T = 1$, or an average over a long term for $T > 1$. The former is useful to study the short term investment, while the latter could be attractive for the longer term. When a stationary process is assumed for $y_t^{(\tau)}$, $\mathbb{E}\left(\bar{y}\right)$ is the same for both $T = 1$ and $T > 1$. However, the impact of model uncertainty is not necessarily the same. This will be discussed in Section 2.2.3.

Consider $\bar{y} = V'y$ where $V = \frac{1}{T}\iota$ with $\iota \in \mathbb{R}^{T\times 1}$ a vector of ones. For a nominal ATSM, y follows a joint multivariate normal distribution with mean $\mu_{NOM} \in \mathbb{R}^{T\times 1}$ and covariance matrix $\Sigma_{NOM} \in \mathbb{R}^{T\times T}$, i.e., $y \sim \mathcal{N}\left(\mu_y, \Sigma_y\right)$ under $p_{NOM}(y)$. After the change of measure by m_θ^*, y is also normally distributed , i.e., $y \sim \mathcal{N}\left(\mu_y + c\theta\Sigma_y V, \Sigma_y\right)$ under $p_{\theta,NOM}(y)$; see Appendix 2.A for the derivation details. Thus, the bound expectations of $\bar{y}$ can be expressed alternatively by

$$\mathbb{E}_j(\bar{y}) = \mathbb{E}_j(V'y) = \mathbb{E}_{NOM}\left[V'\left(\mu_y + c\theta\Sigma_y V\right)\right] = \mathbb{E}_{c\theta}(\bar{y}). \tag{2.16}$$

More specifically, the misspecification interval of $\mathbb{E}_{NOM}(\bar{y})$ is given by

$$MI(\bar{y}) = \left[V'\left(\mu_y + \theta\Sigma_y V\right), V'\left(\mu_y - \theta\Sigma_y V\right)\right]. \tag{2.17}$$

Applying (2.9), the KL divergence between these two multivariate normal distributions yields

$$D\left[p_{\theta,NOM}(y)||p_{NOM}(y)\right] = \kappa_\theta = \frac{1}{2}\theta^2 V'\Sigma_y V, \tag{2.18}$$

θ^* follows straightforwardly by the condition $\kappa_{\theta^*} = \kappa^*$, and determines $\mathbb{E}_{c\theta^*}(\bar{y})$.

So far, with a determined κ^*, θ^* can be obtained by either (2.15) or (2.18). The former applies for general cases, regardless of the types of distributions, while the latter is specific for two normal distributions with the same variance. The bounds of $MI(\bar{y})$ for κ^*, $\mathbb{E}_{c\theta^*}(\bar{y})$, can also be evaluated by either (2.10) or (2.16).

The Impact of Model Uncertainty and the Dynamic Process Lengths

In formula (2.9), the divergence κ^* typically increases in T, the length of the process. Moreover, given $\kappa_\theta = \kappa^*$, κ_{θ^*} also increases in T. From (2.18), this can be easily verified that θ increases in κ_θ in case of two normal distributions with the same variance. In other words, θ^* increases in T. Eventually, the increased θ^* affects the the upper and the lower bound expectations for $\mathbb{E}_{NOM}(\bar{y})$. Appendix 2.B proves that the misspecification interval is monotonically widened in θ^*, for $\theta^* > 0$. Together this leads to a general claim.

Claim 1. *Suppose y_t is a stationary process of length T, and the autocorrelation of y_t is positive. For a function $g : \mathbb{R}^T \to \mathbb{R}$ of a multivariate variable y, the misspecification interval for $\mathbb{E}_{NOM}[g(y)]$ becomes wider for longer periods T.*

The essential reason behind this relation is the higher dimension of the joint distribution of y, associated with T. The misspecification interval is affected by the full length of the process. The longer the process is, the stronger the impact is, and the wider the misspecification interval is. In return, a wider misspecification interval implies that the decision based on this impact is inclined to be more conservative. The least conservative decision would be based on the impact for $T = 1$.

Moreover, the optimization problem in (2.11) implies that both $\mathbb{E}_{NOM}(\bar{y})$ and $\mathbb{E}_*(\bar{y})$ should lie within the misspecification interval. Since the quantities of $\mathbb{E}_{NOM}(\bar{y})$ and $\mathbb{E}_*(\bar{y})$ are not related to T, but the impact of model uncertainty is, and increases in T, one may set $T = 1$ for a primary investigation. If the inclusion situation is verified for $T = 1$, it will be satisfied for larger T. For this purpose, in the empirical analysis I will focus on $T = 1$.

2.3 Empirical Analysis

This section implements the empirical investigation on the impact of model uncertainty in the context of a class of ATSMs. The data are described first. It is followed by the implementation of the MCS selection procedure to identify a set of indistinguishably best-performing models. Then I will quantify the amount of model uncertainty, and construct the uncertainty set, by taking the MCS as an approximate location of the true DGP. Next, the uncertainty parameters and misspecification intervals are obtained. In this empirical analysis, I compare the results for several chosen nominal models to price bonds with some different time-to-maturities. Furthermore, to illustrate the impact of model uncertainty transmitted and affected on other financial products, I investigate a max-10-year annuity, using the modelled bond yields in discounting.

2.3.1 Data Description

The factors used in this paper are individual bond yields with different time-to-maturities τ, given by the U.S. treasury securities which are available only for limited time-to-maturities. When the information for other time-to-maturities is needed, a convenient way to get it is to apply the Nelson-Siegel-Svensson (NSS) approach [67] to estimate it by the NSS parameters. Gürkaynak et al. [30] publish daily their NSS parameters estimates for the U.S. treasury securities, starting from June 1961. The bond yields given by these parameters are continuously compounded and effective annualized rates. In this paper, I will use these parameters in monthly frequency, from June 1961 to December 2016, to obtain the bond yields with any needed time-to-maturities. For each bond yield, I have 666 observations. The time range starts from the beginning period in the data to the recent date, including as much information as possible for the purpose of reducing parameter uncertainty.

Table 2.C.1 reports the basic statistical descriptions of the NSS yield estimates for the time-to-maturity of 10, 30, 40, 50, 60, 80, 90, and 110 months. These maturities are chosen *ad hoc*. The means of the bond yields suggests an increasing yield curve from 5.14% to 6.27%. The bond yields with longer time-to-maturity fluctu-

ate less. Figure 2.C.1 depicts the dynamics of the estimated bond yields. Generally, they feature with an upward trend until the early 80's, then the yields decreases to the historical lows in the recent years.

2.3.2 The MCS Selection

The MCS selection starts with an initial collection of potentially empirically relevant models. This initial collection could include any model a practitioner has in mind. They can be different in terms of assumptions, estimations, structures, variables, etc. This paper only considers models with a linear structure, different only in terms of the factor choices. In general, the initial collection should include as many types of models as possible, regardless of how many or which factors to choose.[8] I start from

$$\begin{aligned}\mathcal{M}_0 =\{&[1], [6], [9], [24], [60], [84], [120], [1, 6], [6, 9], [9, 12], [12, 36], [12, 60],\\ &[36, 84], [60, 120], [1, 6, 9], [6, 9, 12], [12, 24, 60], [36, 60, 84], [12, 60, 120],\\ &[1, 6, 9, 12], [6, 9, 12, 24], [24, 36, 60, 84], [6, 9, 12, 24, 36], [6, 24, 36, 60, 84],\\ &[1, 6, 9, 12, 24], [6, 12, 24, 36, 60, 120], [9, 12, 36, 60, 84, 120],\\ &[1, 6, 9, 12, 24, 36, 60], [12, 24, 36, 60, 84, 120], [6, 9, 12, 24, 36, 60, 84, 120]\},\end{aligned}$$

which is a set including 30 vector elements to represent the models for the MCS selection. The factors of these models are the bond yields with the maturity of months as indicated in the vectors. The models for the MSC selection include one-factor to eight-factor models, and maturities 1-month to 120-month yield. These models are used to price the bonds with the maturities 10-month, 30-month, 40-month, 50-month, 60-month, 80-month, 90-month, and 110-month, the maturities chosen to report in Table 2.C.1. I set the significance level at $\alpha = 5\%$, which is a commonly used one.

I apply the 3-step OLS approach by Adrian et al. [1] to estimate these models.

[8] This is learned from some experiments that I have implemented. These experiments try several initial collections with a number of models using a variety of combinations of factors. Details of these experiments are available on request.

The eigenvalues of $\widehat{\Upsilon}$ for all the models are not larger than 1. In other words, the factors are stationary processes, and thus so are the yield processes.

Figure 2.C.2 provides a general picture of the performance of the models in $\mathcal{M}_0$ for the chosen maturities, as compared to the NSS yield estimates. In the picture, the horizontal axis is the time line while the vertical axis is the bond yields in percentage. Each dashed line shows the estimates by each model, while the solid line shows the NSS estimates. Generally speaking, the model using a single factor of 1-month yield behaves poorly the most. The longer the maturity is, the worse it performs. The other models do not deviate from the NSS estimates as much as the model with one factor of 1-month yield, and fit the NSS estimates much better. But when pricing longer term bonds, their estimates are relatively more dispersed. Intuitively, this means that the MCS procedure could distinguish them much easier, and result in less remaining models in the MCS.

The models in $\mathcal{M}_0$ are firstly ranked according to the values of the loss function in (2). The rankings are reported in Table 2.C.2a for each τ-month bonds. In this table, the first column assigns index j to each model in $\mathcal{M}_0$. The second column show the specific factors. For each τ-month-maturity, the models' rankings are reported in descending order, according to the values of the loss function. Some presumptive nominal models could be chosen from $\mathcal{M}_0$, as indicated by an asterisk "$*$". Their rankings for performance are marked as well.

The rankings show that the worst models mainly use one factor, and the best ones use a number of factors. A model with factors of short rates perform poorly for pricing a long term bond, and vice versa. For instance, the rankings confirm that the model using a 1-month yield factor is the worst. Model 30 using 8 factors from short rates to long rates performs the best, especially in pricing medium and long bonds.

The MCS elimination procedure is implemented by applying the moving block bootstrap; see Goncalves and White [28]. The moving block bootstrap approach is able to capture the possible serial dependence. How well it could work relies on the chosen window length L. Too short a length might fail to consider the serial dependence, while too long might reduce the randomness of resampling. I have carried out the experiments for the window lengths of $L = 50, 100$, and 200 periods. These lengths result in different sizes of $\mathcal{M}^*_{95\%}$, summarized in Table 2.C.2c. Generally speaking, the $\mathcal{M}^*_{95\%}$ size significantly shrinks as L is widened until 100 periods. The comparison shows that $L = 100$ provides the smallest $\mathcal{M}^*_{95\%}$, a signal that the data are used most informatively. Thus, $L = 100$ is chosen for further analysis.

Table 2.C.2b reports the p_r-val values for each round of test against the models one by one from top to bottom. The remaining elements in the $\mathcal{M}^*_{95\%}$, the MCS with 95% confidence level, are marked in bold in the table as well as in Table 2.C.2a. The other elements are eliminated and no longer considered. The largest $\mathcal{M}^*_{95\%}$ appears in pricing the 60-month bond, containing 15 models. Smaller $\mathcal{M}^*_{95\%}$ are in pricing the 30-, and the 40-month bonds. For the maturities of 80 months and 90 months, the data is informative enough to identify a unique best-performing model. When a model falls into the $\mathcal{M}^*_{95\%}$, for instance, Model 25 in pricing 10-month bond, it is regarded as statistically indistinguishable from the true DGP. In other words, only parameter uncertainty exist between them. For the others falling outside of the $\mathcal{M}^*_{95\%}$, there exist model misspecification uncertainty as well.

2.3.3 The Empirical Impact of Model Uncertainty

The MCS collects the indistinguishable best-performing model(s), deemed to represent approximately the true DGP given a finite sample. When a nominal model lies outside of the MCS, extra model misspecification uncertainty is introduced. In order to capture the true DGP, the uncertainty set $\mathcal{P}_U$ is determined by the KL divergence taking into account the MCS. The impact of model uncertainty is evaluated within $\mathcal{P}_U$, resulting in the misspecification interval of a nominal model.

The remainder of this section will firstly choose several nominal models to investigate. It is followed by discussing the simulation method used in evaluating the impact of model uncertainty. Next, I determine κ^* by the KL divergence formula in (2.9). This allows me to obtain the uncertainty parameter θ^* by either (2.15) or (2.18). At last, with θ^*, I obtain the misspecification intervals for the chosen nominal models, and compare them.

The Chosen Nominal Models

Several nominal models are chosen to investigate for comparison. They are different from the factor choice X_{NOM}. These models are the seven nominal models indicated in Table 2.C.2a, namely,

$$NOM_1 = [1]; \quad NOM_2 = [120]; \quad NOM_3 = [12, 36]; \quad NOM_4 = [12, 60];$$
$$NOM_5 = [6, 9, 12]; \quad NOM_6 = [12, 60, 120]; \quad NOM_7 = [1, 6, 9, 12, 24].$$

NOM_1 is a single-factor model using the 1-month short rate as the single factor. It performs the worst across all maturities, shown in Table 2.C.2a. This single factor model can be seen as a representative of Vasicek [69]'s model, and it is used widely in practice; see Brigo and Mercurio [6]. NOM_2 uses a single-factor of the long rate of the 120-month that performs better for longer maturities. NOM_3 and NOM_4 are two-factor models with a common 12-month short rate, but with a different medium rate (60-month v.s. 36-month). NOM_5 and NOM_6 are three-factor models. NOM_5 uses only short rates. NOM_6 includes a short rate, a medium rate, and a long rate, echoing the canonical Litterman and Scheinkman [52]'s three-factor model. NOM_7 is a five-factor model that has significant divergences from the $\mathcal{M}^*_{95\%}$ for most of the maturities. In Table 2.C.2a, NOM_7 in pricing 10-month bond, and NOM_4 and NOM_6 in pricing 60-month bond are in the corresponding $\mathcal{M}^*_{95\%}$. This implies that they have only parameter uncertainty when pricing the bonds with those maturities.

Simulation and the Determination of κ^*

I determine κ^* is by the formula in (2.9), which depends on the means and variances of the bond yields. These bond yields are obtained by simulation. Specifically, an ATSM is characterized by A_τ, B_τ, σ_u, μ, Φ, and σ_v. As these characterizing parameters are estimated respectively by the 3-OLS approach, I apply the Monte-Carlo simulation approach for sampling the conditional bond yield $y^{(\tau)}$. The initial value X_0 is drawn from to the multivariate normal distribution with the historical mean and variance of the factor vector. A simulation of an ATSM bond yield $y^{(\tau)}$ consists of N paths of a T-period process. In this paper, I simulate $N = 2,000,000$ paths for each simulation, and $T = 1$ for the purpose discussed earlier in Section 2.2.3. The largest divergence between a nominal model and the MCS ones is considered as $\widehat{\kappa}^*$. To improve precision, I implement $N_0 = 150$ times simulations to estimate $\widehat{\kappa}^*$, and use the expectation given by these simulations.

Table 2.C.3 collects the expectations and standard deviations of the $\widehat{\kappa}^*$, reported in basis points. The bold and underlined items signal the divergences significantly different from 0, at 95% confidence interval. It can be seen that the misspecification uncertainty quantified by $\widehat{\kappa}^*$ is remarkably the largest when using NOM_1 of a single 1-month bond yield factor. Noticeable is NOM_6, which is a three-factor model using a short rate, a medium rate and a long rate. The misspecification uncertainty is too little to be different from 0 across all maturities, meaning that NOM_6 performs almost indifferent from the MCS models. For other nominal models, the misspecification uncertainty is significant when pricing bonds with longer maturities. At last, if compared with Table 2.C.2a, one may find nominal models falling outside of the MCSs could have divergences indifferent from 0. This might results from the errors in moving block bootstrap, simulation, and estimation, particularly when the differences between models are quite small.

By the simulation sample of a bond yield under a nominal model, we are also able to estimate the probability distribution $p_{NOM}(y)$, i.e., $y^{(\tau)} \sim \mathcal{N}\left(\widehat{\mu}_y, \widehat{\Sigma}_y\right)$. Then, the $y^{(\tau)}$ ruled by an alternative model $p_{\theta,NOM}(y)$ is $y^{(\tau)} \sim \mathcal{N}\left(\widehat{\mu}_y + c\theta\widehat{\Sigma}_y V, \widehat{\Sigma}_y\right)$. The uncertainty parameter θ determined by (2.18) depend on Σ_y. Therefore, I reported the estimates $\widehat{\Sigma}_y$ in Table 2.C.4. We can see that the $\widehat{\Sigma}_y$ for NOM_1 are generally the smallest. Together with the largest $\widehat{\kappa}^*$, we expect that the corresponding estimates of θ^* would be much larger than those using other nominal models.

The Misspecification Intervals

The bounds of a misspecification interval can be given by either (2.10) or (2.16). These two approaches provide the same empirical results. Following (2.17), the main part determining the width of misspecification interval $\theta V' \Sigma_y V$ requires θ, besides Σ_y. The estimates $\widehat{\theta}^*$ corresponding to the $\widehat{\kappa}^*$ in Table 2.C.3 are reported in the Panel A of Table 2.C.5. Considering that the divergence should be non-negative, and the negative $\widehat{\kappa}^*$ in Table 2.C.3 are indifferent from 0, I replace these negative items to be 0 when estimating θ^*, resulting in $\widehat{\theta}^* = 0$ items in this sub-table.

Table 2.C.3 also reports the upper and lower bounds of the misspecification intervals in Panel D and Panel E, respectively. Panel B and Panel C are the the bond yields estimated by the nominal models and the MCS ones, which are expected to lie within the misspecification intervals. Figure 2.C.3 plots these intervals across maturities, which form the bound yield curves. The yield curves of the nominal models and the MCS models are plotted as well. One can see that the nominal and the MCS yield curves are generally within the bound yield curves, satisfying the inclusion situation for $T = 1$. When $T > 1$, the misspecification interval of $\mathbb{E}_{NOM}(\bar{y})$ will be wider, and the inclusion would be more obvious.

The size of a misspecification interval reflects the impact of misspecification uncertainty. I compare the width of one side of the interval, $\theta V' \Sigma_y V$, to the bond yield under $p_{NOM}(y)$, $\mathbb{E}_{NOM}(\bar{y})$. The ratio provides with the information how far the true expectation could deviate from $\mathbb{E}_{NOM}(\bar{y})$ due to misspecification uncertainty. Panel F in Table 2.C.5 reports these ratios. It shows that the misspecification uncertainty could cause a change up to 16.11% of $\mathbb{E}_{NOM}(\bar{y})$. For those whose misspecification uncertainty is significantly different from 0, the lowest ratio is 0.66%. These ratios would be higher when $T > 1$, suggesting non-negligible impact.

Next, I particularly address several nominal model cases. The first one is NOM_1. The misspecification interval gets narrower in maturity. The essential reason is that the corresponding $\widehat{\kappa}^*$ is decreasing. Secondly, for NOM_3, NOM_5 and NOM_7, the misspecification intervals are substantially wider for pricing longer bonds. This is easy to understand, because these models use factors of short rates, and thus the misspecification uncertainty is larger for longer maturities. Thirdly, NOM_2, NOM_4, and NOM_6 show little change because of limited misspecification uncertainty. This means that the decision maker would suffer much less from the model uncertainty if using these models in practice.

Last but not least, although $\mathbb{E}_{NOM}(\bar{y})$ and $\mathbb{E}_*(\bar{y})$ have to lie within the interval, one is not necessary higher than the other. This is because $\mathbb{E}_*(\bar{y})$ is given by the model fitting the data the best, with no guarantee of a higher value than $\mathbb{E}_{NOM}(\bar{y})$. This can be seen in Figure 2.C.3 when using NOM_1 to price longer maturity bonds. Observing their values is helpful in understanding the conservativeness in decision-making. Use as an example the case of NOM_1 for a long maturity. A conservative decision maker may consider the lower bound as the worst case; for instance, a lender would worry about low interest rate, and hence the lower bound is the worst case for him to pay attention to. He may also consider the upper bound as the worst case; for instance, a borrower worries about high interest rate. Then, the decision made in the former situation could be less conservative than that in the latter situation, because the $\mathbb{E}_*(\bar{y})$ of the best-performing model is closer to the lower bound.

2.3.4 Further Impact on An Annuity

ATSMs are an essential tool to determine the interest rates, and interest rates are widely used in finance. Therefore, the impact of model uncertainty on yields can be transmitted further to affect other financial products. As a simple application, an annuity, using yields for discounting, is illustrated to demonstrate how the impact of model uncertainty transmitted from the bond yields.

In the previous investigation, the yields are studied for eight maturities from 10 months to 110 months. The yields for other maturities are obtained by linear interpolation or extrapolation. To avoid too much errors, I only consider an annuity given for maximum 10 years (120 months).

The annuity pays a fixed amount of 1 Euro to the annuitant every year as long as he is alive. In life insurance, the annuity is also affected by the survival rate. Denote by ${}_{\tau}p_x$ with $\tau = 0, \cdots, 10$ the probability of surviving τ year given age x. The present value of the annuity for a person of age x is discounted by the expected interest rate $\bar{y}_t^{(\tau)}$, given by

$$A_x = \sum_{\tau=0}^{10} {}_{\tau}p_{x,t} \cdot \exp\left(-\tau \bar{y}_t^{(\tau)}\right).$$

The survival rate is calculated based on the published expected Dutch mortality rates.[9] The model uncertainty of these estimated mortality rates is neglected.

Worthy to note is that, this illustration does not look into the probability distribution of annuity affected by the bond yields, so as to investigate the impact of model uncertainty. Although this may sound like a more reliable approach, it could be more complicated. Secondly, because of the specific structure of the annuity formula, investigating from the perspective of the annuity's distribution may result in unbounded impact. The solution approach to this problem could be the topic of future research. For these reasons, I only apply the resulting impact of model uncertainty on bond yields, and study the transmitted impact of model uncertainty from the bonds. As a consequence, the misspecification intervals of the annuity are determined by the bounds of the misspecification intervals of the bond yields.

I investigate the annuity for the present age $Age = [25, 45, 65, 85]$, for males and for females. The results are reported in Table 2.C.6. In the table, the upper and lower bounds of the misspecification intervals are reported as the differences to the expected annuity under the nominal models, with the symbols *UB* and *LB*.

Firstly, it can be seen from the table that, the largest misspecification interval is when using NOM_1, because of the most significant impact of model uncertainty on bond yields. One side of the interval could be as large as 0.22 Euro cents. The

[9] The data are published in 2016, by the Royal Actuarial Society of the Netherlands (http://www.ag-ai.nl/). This information includes the mortality rate for a person from age 0 to age 120, from the present year to the next 120 years. It reports the mortality rates for males and for females, respectively. In this application, only the data of the next 10 successive years of the investigated ages are taken.

second largest interval is given by NOM_5, whose overall intervals for bond yields is also second largest to those of NOM_1. The model uncertainty using NOM_6 has almost no impact on the annuity. Generally, if compared with the misspecification intervals of the bond yields, those of the annuity are much small, resulting from the discounting effect.

Secondly, it can also be observed that the misspecification interval shrinks in age, because of the decreasing survival rate. Moreover, the annuity of male is typically lower than that of female, because of the lower survival rate of the male. For the same reason, the intervals for female are slightly wider than those for male.

At last, these intervals of annuity are subjected to $T = 1$ bond yield processes, and paying 1 euro per year. When assigning much larger values to these two conditions, the impact of model uncertainty would become non-negligible.

2.4 Summary and Conclusions

ATSMs play a significant role in explaining, analyzing and predicting bond yields by a linear regression on factors. However, some simple factors preferred by practitioners might inevitably result in significantly discrepant models from the true DGP. The discrepancy is considered as the model uncertainty of the nominal model from the true DGP. This paper is an empirical study to evaluate the impact of model uncertainty on bonds priced by ATSMs, employing Schneider and Schweizer [65]'s approach. Results are compared between several nominal models different in terms of factors choice. Furthermore, an annuity associated with bond yields is studied to illustrate the impact of model uncertainty transmitted from bond yields modelling.

One finding in this paper is that, when a longer yield process is taken into account, the impact on the expected yield over time is stronger, resulting in a wider misspecification interval and more conservative attitude. This finding is supported by empirical analysis and partly proved in theory. Moreover, the empirical investigation suggests a considerable impact of model uncertainty when using a single-factor model of a 1-month short rate. The magnitude mounts up to 32.22% of the nominal outcomes. Realizing this is important because such nominal models, especially this model, could be broadly used in industrial practice. The least impact of model uncertainty appears when the factors in the nominal model include a short rate, a medium rate, and a long rate. This is aligned with the interpretation of level, slope, and curvature in the canonical three-factor model that fits remarkably well empirically.

This paper also makes a simple analysis of a max-10-year annuity to investigate the impact of model uncertainty on bond yields transmitted to other financial products. Because the yields are used in discounting, the model uncertainty affects an annuity much less significantly. Nevertheless, when applying to large investment projects and expected yields for longer term are considered, the impact of model uncertainty will not be negligible.

In this paper, the investigation on model uncertainty is illustrated within a setting as simple as possible. In essence, the estimation employs the 3-step OLS approach, and the factors in use are individual bond yields. This helps in understanding the mechanism of the approach to evaluate model uncertainty, and the fundamental impact of model uncertainty when using certain nominal models. Based on this investigation, more sophisticated settings can be considered. For instance, future studies can include factors of macro-economic indices, latent variables, principal components, etc. They can also consider more advanced approaches, such as investigating the bond yield dynamics and prediction, whether taking into account arbitrage-free restrictions, other likelihood-based estimations, etc. Further topics can also be extended to financial theories to embed model uncertainty, and to empirical financial studies such as asset-liability analysis built on ATSMs.

Appendix

Appendix 2.A The Distribution of An Alternative Model

Suppose the distribution of a multivariate nominal model of $y = (y_1, \cdots, y_T)' \sim \mathcal{N}\left(\mu_y, \Sigma_y\right)$ is given by

$$p_{NOM}(y) = \frac{1}{\sqrt{(2\pi)^T |\Sigma_y|}} \exp\left[-\frac{1}{2}\left(y - \mu_y\right)' \Sigma_y^{-1}\left(y - \mu_y\right)\right].$$

The quantity of interest depending on y is $\mathbb{E}(\bar{y}) = \mathbb{E}\left(\frac{1}{T}\sum_{t=1}^{T} y_t\right) = \mathbb{E}(V'y)$, where $V = \frac{1}{T}\iota$ and $\iota \in \mathbb{R}^{T\times 1}$. By Proposition 1, the distribution of an alternative model is given by $p_{\theta,NOM}(y) = m_\theta^* \cdot p_{NOM}(y)$, where $m_\theta^* = \frac{\exp(c\theta\bar{y})}{\mathbb{E}[\exp(c\theta\bar{y})]}$, for $c = 1$ indicating the upper bound and $c = -1$ the lower bound. Then, rewriting this distribution gives

$$
\begin{aligned}
p_{\theta,NOM}(y) &= \frac{\exp\left(c\theta\bar{y}\right)}{\mathbb{E}\left[\exp\left(c\theta\bar{y}\right)\right]} \cdot \frac{1}{\sqrt{(2\pi)^T|\Sigma_y|}} \exp\left[-\frac{1}{2}\left(y-\mu_y\right)'\Sigma_y^{-1}\left(y-\mu_y\right)\right] \\
&= \frac{1}{\sqrt{(2\pi)^T|\Sigma_y|} \cdot \mathbb{E}\left[\exp\left(c\theta\bar{y}\right)\right]} \exp\left[-\frac{1}{2}\left(y-\mu_y\right)'\Sigma_y^{-1}\left(y-\mu_y\right) + c\theta V'y\right] \\
&= \frac{1}{\sqrt{(2\pi)^T|\Sigma_y|\mathbb{E}\left[\exp\left(c\theta\bar{y}\right)\right]}} \\
&\cdot \exp\Bigg\{-\frac{1}{2}\Bigg[\left(y-\left(\mu_y+c\theta\Sigma_y V\right)\right)'\Sigma_y^{-1}\left(y-\left(\mu_y+c\theta\Sigma_y V\right)\right) \\
&+ \left(-2c\theta\mu_y - c^2\theta^2\Sigma_y V\right)\Bigg]\Bigg\} \\
&= \frac{1}{\sqrt{(2\pi)^T|\Sigma_y|}} \exp\left\{-\frac{1}{2}\left[\left(y-\left(\mu_y+c\theta\Sigma_y V\right)\right)'\Sigma_y^{-1}\left(y-\left(\mu_y+c\theta\Sigma_y V\right)\right)\right]\right\} \\
&\cdot \frac{\exp\left(c\theta\mu_y + \frac{1}{2}c^2\theta^2\Sigma_y V\right)}{\mathbb{E}\left[\exp\left(c\theta\bar{y}\right)\right]}.
\end{aligned}
\tag{2.19}
$$

By the moment generating function of normal distribution, $\mathbb{E}\left[\exp\left(c\theta\bar{y}\right)\right] = MGF_{\bar{y}}\left(c\theta\right) = \exp\left(c\theta\mu_y + \frac{1}{2}c^2\theta^2\Sigma_y V\right)$. Therefore, (2.19) is simplified as

$$
p_{\theta,NOM}(y) = \frac{1}{\sqrt{(2\pi)^T|\Sigma_y|}} \exp\left\{-\frac{1}{2}\left[\left(y-\left(\mu_y+c\theta\Sigma_y V\right)\right)'\Sigma_y^{-1}\left(y-\left(\mu_y+c\theta\Sigma_y V\right)\right)\right]\right\}.
$$

In other words, $y \sim \mathcal{N}\left(\mu_y + c\theta\Sigma_y V, \Sigma_y\right)$, whose mean equals to the nominal mean of y plus a constant vector $c\theta\Sigma_y V$, and variance remains the same.

Appendix 2.B The Misspecification Interval Increases in θ

The misspecification interval is the simple distance between the expectations of the bound models of the uncertainty set for a nominal model, subject to the misspecification uncertainty amount κ.

Theorem 1. *Assuming Gaussian distribution, the misspecification interval increases in θ, because the first order condition equal to $cVar_{c\theta}(\bar{y})$ is larger than 0 for $c = 1$ and smaller 0 for $c = -1$, where $Var_{c\theta}(\bar{y})$ is the variance of $\bar{y}$ under $p_{\theta,NOM}(y)$.*

Proof. The bounds $\mathbb{E}_{c\theta}(\bar{y})$ can be obtained by (2.10) or (2.16). If $\mathbb{E}_{c\theta}(y) = \mathbb{E}_{NOM}\left(m_\theta^* \cdot y\right)$ in (2.10) and m_θ^* in (2.14), we have $\mathbb{E}_{NOM}\left(m_\theta^* \cdot \bar{y}\right) = \mathbb{E}_{NOM}\left\{\frac{\exp(c\theta\bar{y})\bar{y}}{\mathbb{E}_{NOM}[\exp(c\theta\bar{y})]}\right\}$. The first derivative with respect of θ is

$$
\begin{aligned}
&\frac{\partial}{\partial\theta}\mathbb{E}_{NOM}\left(m_\theta^* \cdot \bar{y}\right)\\
=&c \cdot \frac{\mathbb{E}_{NOM}\left[\exp\left(c\theta\bar{y}\right)\bar{y}^2\right]\mathbb{E}_{NOM}\left[\exp\left(c\theta\bar{y}\right)\right] - \mathbb{E}^2_{NOM}\left[\exp\left(c\theta\bar{y}\right)\bar{y}\right]}{\mathbb{E}^2_{NOM}\left[\exp\left(c\theta\bar{y}\right)\right]}\\
=&c \cdot \frac{\mathbb{E}_{NOM}\left[\exp\left(c\theta\bar{y}\right)\bar{y}^2\right]}{\mathbb{E}_{NOM}\left[\exp\left(c\theta\bar{y}\right)\right]} \cdot \frac{\mathbb{E}_{NOM}\left[\exp\left(c\theta\bar{y}\right)\cdot 1\right]}{\mathbb{E}_{NOM}\left[\exp\left(c\theta\bar{y}\right)\right]} - c \cdot \left\{\frac{\mathbb{E}_{NOM}\left[\exp\left(c\theta\bar{y}\right)\bar{y}\right]}{\mathbb{E}_{NOM}\left[\exp\left(c\theta\bar{y}\right)\right]}\right\}^2\\
=&c \cdot \mathbb{E}_{c\theta}(\bar{y}^2)\mathbb{E}_{c\theta}\left(1\right) - c \cdot \left[\mathbb{E}_{c\theta}(\bar{y})\right]^2\\
=&c \cdot \mathbb{E}_{c\theta}(\bar{y}^2) - c \cdot \left[\mathbb{E}_{c\theta}(\bar{y})\right]^2\\
=&c \cdot Var_{c\theta}(\bar{y}),
\end{aligned}
$$

Since y has a non-zero variance Σ_y, $Var_{c\theta}(\bar{y})$ is also larger than 0. Then, the result of the first derivative suggests that $\mathbb{E}_{c\theta^*}(\bar{y})$ monotonically increases in θ for $c = 1$ and decreases for $c = -1$. Hence, the misspecification interval increases in θ.

If using (2.16), we have $\mathbb{E}_{c\theta}(\bar{y}) = \mathbb{E}_{NOM}\left[V'\left(\mu_y + c\theta\Sigma_y V\right)\right] = \mu_y + c\theta V'\Sigma_y V$. The monotonic increasing relation between the misspecification interval and θ is trivial. □

Appendix 2.C Tables and Figures

Table 2.C.1: Data descriptions of the NSS estimates for the annualized bond yields with selected time-to-maturities. The first column indicates the maturities of τ months. The second and the third columns are the means and standard deviations. The forth and fifth columns report the ranges of the NSS estimates.

τ(month)	mean(%)	std(%)	minimum(%)	maximum(%)
10	5.14	3.30	0.09	15.84
30	5.51	3.16	0.27	15.28
40	5.65	3.08	0.38	15.06
50	5.77	3.01	0.51	14.86
60	5.88	2.95	0.64	14.69
80	6.07	2.83	0.94	14.44
90	6.14	2.79	1.09	14.34
110	6.27	2.71	1.39	14.20

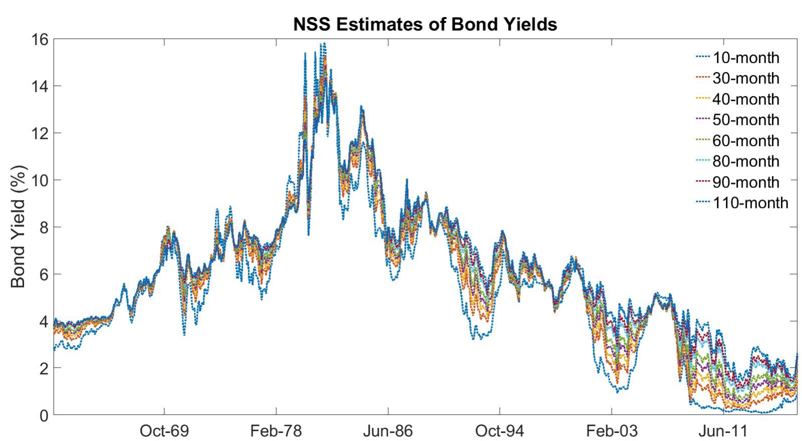


Figure 2.C.1: NSS estimates of bond yields with maturities of 10, 30, 40, 50, 60, 80, 90, and 110 months, ranging from June 1961 to December 2016.

Table 2.C.2: MCS selection of $\mathcal{M}_0$ at the significance level $\alpha = 5\%$. Table 2.C.2a shows the selection of the MCS, using a 100-period window length in bootstrapping. Boldfaced models are included in the corresponding MCSs $\mathcal{M}^*_{95\%}$. Table 2.C.2b collects the p_r-val values when testing in each round against the corresponding model listed in Table 2.C.2a; the sizes of the MCSs are summarized in the last row. Table 2.C.2c compares the MCS sizes obtained using different window lengths of the moving block.

(a) The first column assigns index to the models in $\mathcal{M}_0$ listed in the second column. In each remaining column are model indexes for pricing the bond with τ-month time-to-maturity, ranked in descending order according to their values of the loss function. In each of these columns, the model in the first row performs the worst, while the one in the last row performs the best. The asterisk $*$ indicates the possible nominal models.

j	X_{NOM}	τ-month							
		10	30	40	50	60	80	90	110
1	$[1]^*(NOM_1)$	1*	1*	1*	1*	1*	1*	1*	1*
2	[6]	7*	7*	2	2	2	2	2	2
3	[9]	6	2	8	8	8	8	8	8
4	[24]	5	8	7*	3	3	3	3	3
5	[60]	14	6	3	7*	9	9	9	9
6	[84]	13	3	6	9	15	15	15	15
7	$[120]^*(NOM_2)$	4	5	9	15	10	10	10	10
8	[1,6]	18	9	15	10	4	4	4	16*
9	[6,9]	2	15	10	4	16*	16*	16*	4
10	[9,12]	8	14	16*	16*	7*	20	20	20
11	$[12, 36]^*(NOM_3)$	22	10	4	20	20	21	21	21
12	$[12, 60]^*(NOM_4)$	3	16*	20	6	6	25*	25*	25*
13	[36,84]	12*	20	5	21	21	11*	11*	11*
14	[60,120]	19*	12*	14	25*	25*	7*	5	5
15	[1,6,9]	11*	4	12*	5	11*	5	7*	23
16	$[6, 9, 12]^*(NOM_5)$	17	19*	21	11*	**23**	23	23	12*

To be continued

Continued

j	X_{NOM}	τ-month							
		10	30	40	50	60	80	90	110
17	[12,24,60]	24	13	25*	12*	**13**	12*	12*	17
18	[36,60,84]	29	18	19*	14	**14**	17	17	6
19	$[12, 60, 120]^*(NOM_6)$	10	11*	13	13	**12***	6	6	7*
20	[1,6,9,12]	9	17	11*	19*	**19***	14	14	13
21	[6,9,12,24]	16*	21	17	17	**17**	19*	28	28
22	[24,36,60,84]	30	25*	18	23	**18**	13	19*	18
23	[6,9,12,24,36]	15	**22**	**22**	18	**22**	28	13	14
24	[6,24,36,60,84]	27	**27**	**23**	22	**29**	18	18	19*
25	$[1, 6, 9, 12, 24]^*(NOM_7)$	23	**24**	**24**	**29**	**5**	26	26	22
26	[6,12,24,36,60,120]	26	**23**	**27**	**28**	**24**	22	22	24
27	[9,12,36,60,84,120]	21	**29**	**29**	**24**	**27**	29	29	26
28	[1,6,9,12,24,36,60]	**20**	**30**	**26**	**27**	**28**	24	24	29
29	[12,24,36,60,84,120]	**25***	**26**	**28**	**26**	**26**	27	27	27
30	[6,9,12,24,36,60,84,120]	**28**	**28**	**30**	**30**	**30**	**30**	**30**	**30**

(b) p_r-val values for the moving block window length of 100 periods

Test Round	τ-month							
	10	30	40	50	60	80	90	110
1	0.04%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
2	0.05%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
3	0.06%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
4	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
5	0.09%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%
6	0.16%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%
7	0.16%	0.00%	0.00%	0.00%	0.08%	0.00%	0.00%	0.00%
8	0.23%	0.00%	0.00%	0.00%	0.08%	0.00%	0.00%	0.00%
9	0.24%	0.00%	0.00%	0.00%	0.08%	0.00%	0.00%	0.00%
10	0.26%	0.00%	0.00%	0.00%	0.10%	0.00%	0.00%	0.00%
11	0.32%	0.00%	0.00%	0.00%	0.14%	0.00%	0.00%	0.00%
12	0.45%	0.00%	0.00%	0.00%	1.56%	0.00%	0.00%	0.00%
13	0.47%	0.00%	0.00%	0.00%	2.20%	0.00%	0.00%	0.00%
14	0.55%	0.00%	0.00%	0.00%	2.29%	0.00%	0.00%	0.00%
15	0.60%	0.00%	0.00%	0.00%	3.25%	0.00%	0.00%	0.00%
16	0.95%	0.00%	0.00%	0.00%	**32.11%**	0.00%	0.00%	0.00%
17	0.94%	0.00%	0.00%	0.00%	**36.73%**	0.00%	0.00%	0.00%
18	1.23%	0.00%	0.00%	0.00%	**66.64%**	0.00%	0.00%	0.00%
19	1.25%	0.01%	0.00%	0.00%	**82.03%**	0.00%	0.00%	0.00%

To be Continued

Continued

Test Round	τ-month							
	10	30	40	50	60	80	90	110
20	1.64%	1.02%	0.00%	0.00%	**86.12%**	0.00%	0.00%	0.00%
21	2.41%	1.05%	0.00%	0.00%	**90.50%**	0.02%	0.00%	0.00%
22	2.35%	1.13%	0.10%	0.00%	**93.93%**	0.02%	0.00%	0.00%
23	0.00%	**8.42%**	**13.68%**	0.01%	**97.87%**	0.02%	0.01%	0.00%
24	0.00%	**31.39%**	**27.82%**	0.97%	**98.00%**	0.42%	0.29%	0.01%
25	0.00%	**29.61%**	**66.77%**	**12.56%**	**99.57%**	0.56%	0.04%	0.01%
26	0.00%	**29.67%**	**75.43%**	**12.61%**	**18.41%**	0.62%	0.36%	0.01%
27	0.00%	**33.48%**	**63.46%**	**13.40%**	**19.18%**	0.49%	0.59%	0.40%
28	**10.63%**	**63.96%**	**75.57%**	**22.28%**	**22.19%**	2.00%	1.01%	2.61%
29	**18.73%**	**50.35%**	**81.63%**	**1.53%**	**30.80%**	2.41%	3.33%	**8.22%**
30	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**
$\mathcal{M}^*_{95\%}$ size	3	8	8	6	15	1	1	2

(c) Comparison of the MCS sizes for different moving block window lengths

L (Length)	τ-month							
	10	30	40	50	60	80	90	110
50-period	3	8	8	6	15	1	2	3
100-period	**3**	**8**	**8**	**6**	**15**	**1**	**1**	**2**
200-period	3	8	8	6	18	1	1	1

Table 2.C.3: The divergence $\widehat{\kappa}^*$ calculated with the simulation size of $N = 2,000,000$ and $T = 1$. The expectations and their standard deviations are based on $N_0 = 150$ times simulations. The bold and underlined items are significantly different from 0, at 95% confidence level. The results are reported in basis points in the tables.

j	X_{NOM}	τ-month							
		10	30	40	50	60	80	90	110
		Panel A: $\mathbb{E}(\kappa^*)(\times 10^{-4}$; in basis points)							
1	NOM_1	**<u>468.87</u>**	**<u>628.74</u>**	**<u>428.83</u>**	**<u>291.45</u>**	**<u>200.08</u>**	**<u>106.22</u>**	**<u>79.81</u>**	**<u>46.33</u>**
7	NOM_2	0.36	0.04	0.81	**<u>0.99</u>**	**<u>1.43</u>**	**<u>1.20</u>**	**<u>2.64</u>**	**<u>2.21</u>**
11	NOM_3	0.72	0.72	0.52	0.45	**<u>0.79</u>**	**<u>6.18</u>**	**<u>13.74</u>**	**<u>40.70</u>**
12	NOM_4	**<u>1.16</u>**	1.42	0.06	0.16	0.20	0.32	0.03	**<u>3.20</u>**
16	NOM_5	0.32	0.32	**<u>2.12</u>**	**<u>8.52</u>**	**<u>21.28</u>**	**<u>80.52</u>**	**<u>127.84</u>**	**<u>248.79</u>**
19	NOM_6	0.40	0.55	0.00	0.15	0.14	0.65	0.15	0.00
25	NOM_7	0.37	0.55	0.22	0.60	**<u>2.24</u>**	**<u>14.43</u>**	**<u>28.58</u>**	**<u>73.11</u>**
		Panel B: The standard deviations of $\mathbb{E}(\kappa^*)$ $(\times 10^{-4}$; in basis points)							
1	NOM_1	0.49	0.59	0.55	0.52	0.52	0.51	0.54	0.51
7	NOM_2	0.49	0.45	0.44	0.43	0.44	0.46	0.46	0.47
11	NOM_3	0.41	0.42	0.45	0.40	0.47	0.48	0.45	0.52
12	NOM_4	0.49	0.47	0.50	0.44	0.41	0.49	0.43	0.48
16	NOM_5	0.45	0.47	0.46	0.45	0.46	0.45	0.46	0.52
19	NOM_6	0.45	0.48	0.45	0.49	0.46	0.47	0.48	0.52
25	NOM_7	0.47	0.46	0.46	0.46	0.45	0.47	0.45	0.48

Table 2.C.4: Estimated variances of bond yields, $\widehat{\Sigma}_y$, modelled by the nominal models.

j	X_{NOM}	τ-month							
		10	30	40	50	60	80	90	110
Panel A: Expected sample variances $\widehat{\Sigma}_y$ ($\times 10^{-4}$; in basis points)									
1	NOM_1	7.30	6.28	6.46	6.57	6.63	6.62	6.56	6.49
7	NOM_2	10.91	10.10	9.68	9.29	8.90	8.25	7.98	7.52
11	NOM_3	10.80	9.98	9.51	9.05	8.56	7.66	7.22	6.49
12	NOM_4	10.82	9.95	9.55	9.13	8.74	8.03	7.68	7.12
16	NOM_5	10.95	9.89	9.23	8.57	7.90	6.75	6.25	5.46
19	NOM_6	10.88	9.93	9.49	9.04	8.64	8.02	7.76	7.34
25	NOM_7	10.91	9.99	9.48	8.95	8.44	7.44	7.01	6.22
Panel B: Standard deviations of expected sample variances $\widehat{\Sigma}_y$ ($\times 10^{-8}$)									
1	NOM_1	0.72	0.53	0.56	0.58	0.59	0.59	0.58	0.56
7	NOM_2	1.60	1.37	1.26	1.16	1.06	0.92	0.85	0.76
11	NOM_3	1.57	1.34	1.21	1.10	0.98	0.78	0.70	0.57
12	NOM_4	1.57	1.33	1.22	1.12	1.03	0.86	0.79	0.68
16	NOM_5	1.60	1.31	1.14	0.98	0.84	0.61	0.53	0.40
19	NOM_6	1.58	1.33	1.21	1.10	1.00	0.86	0.81	0.72
25	NOM_7	1.60	1.34	1.20	1.07	0.95	0.75	0.66	0.52

Table 2.C.5: The Impact of Model Uncertainty. The bold and underlined items correspond to those in Panel A, Table 2.C.3.

j	X_{NOM}	τ-month							
		10	30	40	50	60	80	90	110
		Panel A: The estimated θ^*							
1	NOM_1	**11.33**	**14.15**	**11.53**	**9.42**	**7.77**	**5.67**	**4.92**	**3.78**
7	NOM_2	0.26	0.00	0.41	**0.46**	**0.57**	**0.54**	**0.81**	**0.77**
11	NOM_3	0.37	0.00	0.00	0.32	0.43	**1.27**	**1.95**	**3.54**
12	NOM_4	**0.46**	0.00	0.11	0.19	0.00	0.00	0.09	**0.95**
16	NOM_5	0.24	0.00	**0.68**	**1.41**	**2.32**	**4.89**	**6.39**	**9.55**
19	NOM_6	0.27	0.00	0.01	0.00	0.18	0.00	0.20	0.00
25	NOM_7	0.26	0.00	0.00	0.37	**0.73**	**1.97**	**2.86**	**4.84**
		Panel B: The expected yields of the $p_{NOM}(y)$, $E_{NOM}(\bar{y})$							
1	NOM_1	5.14%	5.54%	5.70%	5.83%	5.95%	6.14%	6.22%	6.36%
7	NOM_2	5.16%	5.53%	5.66%	5.78%	5.88%	6.05%	6.12%	6.23%
11	NOM_3	5.15%	5.52%	5.66%	5.78%	5.88%	6.06%	6.13%	6.26%
12	NOM_4	5.15%	5.52%	5.66%	5.78%	5.90%	6.08%	6.15%	6.27%
16	NOM_5	5.14%	5.51%	5.64%	5.77%	5.87%	6.04%	6.12%	6.24%
19	NOM_6	5.14%	5.52%	5.66%	5.78%	5.89%	6.07%	6.15%	6.28%
25	NOM_7	5.14%	5.51%	5.65%	5.77%	5.88%	6.06%	6.13%	6.26%
		Panel C: The expected yields of the $p_*(y)$, $E_*(\bar{y})$							
—	—	5.14%	5.51%	5.65%	5.78%	5.88%	6.07%	6.14%	6.27%

j	X_{NOM}	τ-month							
		10	30	40	50	60	80	90	110
		Panel D: The expected yields of the upper bound models, $E_{\theta}(\bar{y})$							
1	NOM_1	**5.96%**	**6.44%**	**6.44%**	**6.45%**	**6.46%**	**6.52%**	**6.55%**	**6.60%**
7	NOM_2	5.19%	5.53%	5.70%	**5.82%**	**5.93%**	**6.09%**	**6.18%**	**6.29%**
11	NOM_3	5.19%	5.52%	5.66%	5.81%	5.92%	**6.16%**	**6.28%**	**6.49%**
12	NOM_4	**5.20%**	5.52%	5.67%	5.80%	5.89%	6.07%	6.16%	**6.34%**
16	NOM_5	5.17%	5.51%	**5.71%**	**5.89%**	**6.05%**	**6.38%**	**6.52%**	**6.77%**
19	NOM_6	5.17%	5.52%	5.66%	5.78%	5.90%	6.07%	6.16%	6.28%
25	NOM_7	5.17%	5.51%	5.65%	5.80%	**5.94%**	**6.20%**	**6.33%**	**6.56%**
		Panel E: The expected yields of the lower bound model, $E_{-\theta}(\bar{y})$							
1	NOM_1	**4.31%**	**4.66%**	**4.95%**	**5.21%**	**5.43%**	**5.77%**	**5.90%**	**6.11%**
7	NOM_2	5.14%	5.53%	5.62%	**5.74%**	**5.83%**	**6.00%**	**6.05%**	**6.17%**
11	NOM_3	5.11%	5.52%	5.66%	5.75%	5.85%	**5.96%**	**5.99%**	**6.03%**
12	NOM_4	**5.10%**	5.52%	5.65%	5.77%	5.89%	6.07%	6.14%	**6.21%**
16	NOM_5	5.11%	5.51%	**5.58%**	**5.64%**	**5.69%**	**5.72%**	**5.72%**	**5.72%**
19	NOM_6	5.11%	5.52%	5.66%	5.78%	5.87%	6.07%	6.13%	6.28%
25	NOM_7	5.11%	5.51%	5.65%	5.74%	**5.82%**	**5.91%**	**5.93%**	**5.95%**
		Panel F: The ratio $\frac{\theta V' \Sigma_y V}{\mathbb{E}_{NOM}(\bar{y})}$							
1	NOM_1	**16.11%**	**16.02%**	**13.06%**	**10.62%**	**8.66%**	**6.10%**	**5.21%**	**3.85%**
7	NOM_2	0.54%	0.00%	0.70%	**0.74%**	**0.86%**	**0.74%**	**1.06%**	**0.93%**
11	NOM_3	0.77%	0.00%	0.00%	0.49%	**0.62%**	**1.60%**	**2.30%**	**3.67%**
12	NOM_4	**0.97%**	0.00%	0.19%	0.29%	0.00%	0.00%	0.12%	**1.07%**
16	NOM_5	0.51%	0.00%	**1.11%**	**2.10%**	**3.12%**	**5.45%**	**6.54%**	**8.35%**
19	NOM_6	0.58%	0.00%	0.01%	0.00%	0.26%	0.00%	0.25%	0.00%
25	NOM_7	0.55%	0.00%	0.00%	0.57%	**1.04%**	**2.42%**	**3.27%**	**4.83%**

Table 2.C.6: This table collects the misspecification intervals of the max-10-year annuity for male and for female, given by the seven nominal models. The symbol *NOM* refers to the expected annuity of the nominal model. *UB* and *LB* are the *NOM*-centered upper and lower bounds of the interval, respectively.

Age	NOM_1						NOM_2					
	Male			Female			Male			Female		
	LB	NOM	UB	LB	NOM	UB	LB	NOM	UB	LB	NOM	UB
25	-0.0022	**9.9422**	0.0022	-0.0022	**9.9535**	0.0022	0.0003	**9.9428**	0.0003	0.0003	**9.9540**	0.0003
45	-0.0022	**9.8583**	0.0022	-0.0022	**9.8722**	0.0022	0.0003	**9.8588**	0.0003	0.0003	**9.8727**	0.0003
65	-0.0020	**9.1959**	0.0020	-0.0021	**9.4726**	0.0021	0.0002	**9.1964**	0.0002	0.0003	**9.4731**	0.0003
85	-0.0010	**4.9139**	0.0010	-0.0012	**5.6849**	0.0012	0.0001	**4.9141**	0.0001	0.0001	**5.6851**	0.0001
Age	NOM_3						NOM_4					
	Male			Female			Male			Female		
	LB	NOM	UB	LB	NOM	UB	LB	NOM	UB	LB	NOM	UB
25	-0.0008	**9.9427**	0.0008	-0.0008	**9.9539**	0.0008	0.0002	**9.9426**	0.0002	0.0002	**9.9538**	0.0002
45	-0.0007	**9.8587**	0.0007	-0.0007	**9.8726**	0.0007	0.0002	**9.8587**	0.0002	0.0002	**9.8726**	0.0002
65	-0.0007	**9.1963**	0.0007	-0.0007	**9.4730**	0.0007	0.0002	**9.1962**	0.0002	0.0002	**9.4729**	0.0002
85	-0.0002	**4.9141**	0.0002	-0.0003	**5.6851**	0.0003	0.0000	**4.9140**	0.0000	0.0001	**5.6850**	0.0001

Age	NOM_5						NOM_6					
	Male			Female			Male			Female		
	LB	NOM	UB	LB	NOM	UB	LB	NOM	UB	LB	NOM	UB
25	-0.0019	**9.9427**	0.0019	-0.0019	**9.9540**	0.0019	0.0000	**9.9426**	0.0000	0.0000	**9.9538**	0.0000
45	-0.0019	**9.8588**	0.0019	-0.0019	**9.8727**	0.0019	0.0000	**9.8587**	0.0000	0.0000	**9.8726**	0.0000
65	-0.0017	**9.1964**	0.0017	-0.0018	**9.4731**	0.0018	0.0000	**9.1962**	0.0000	0.0000	**9.4729**	0.0000
85	-0.0005	**4.9141**	0.0005	-0.0007	**5.6851**	0.0007	0.0000	**4.9140**	0.0000	0.0000	**5.6851**	0.0000

Age	NOM_7					
	Male			Female		
	LB	NOM	UB	LB	NOM	UB
25	-0.0010	**9.9427**	0.0010	-0.0010	**9.9539**	0.0010
45	-0.0010	**9.8588**	0.0010	-0.0010	**9.8727**	0.0010
65	-0.0009	**9.1963**	0.0009	-0.0009	**9.4730**	0.0009
85	-0.0003	**4.9141**	0.0003	-0.0003	**5.6851**	0.0003

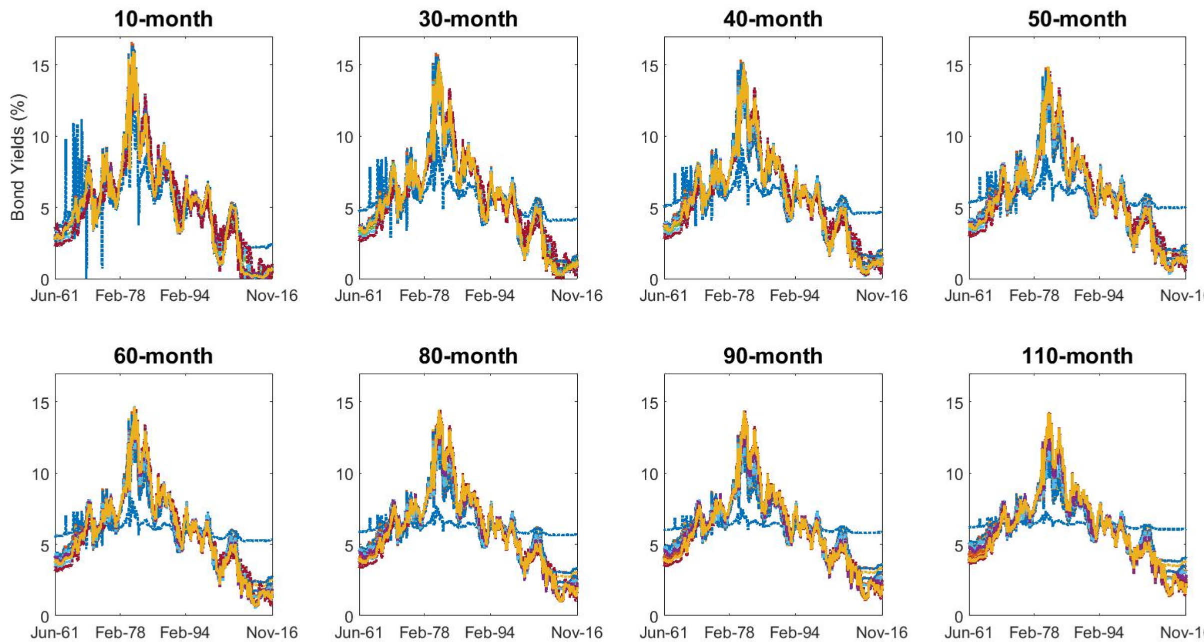


Figure 2.C.2: ATSM estimates of bond yields with various maturities, using the models in $\mathcal{M}_{\text{o}}$ respectively. They are plotted with dashed lines for 666 periods. The solid lines represent the NSS yield estimates. In general, the model using 1-month bond as the factor is the worst-performing one.

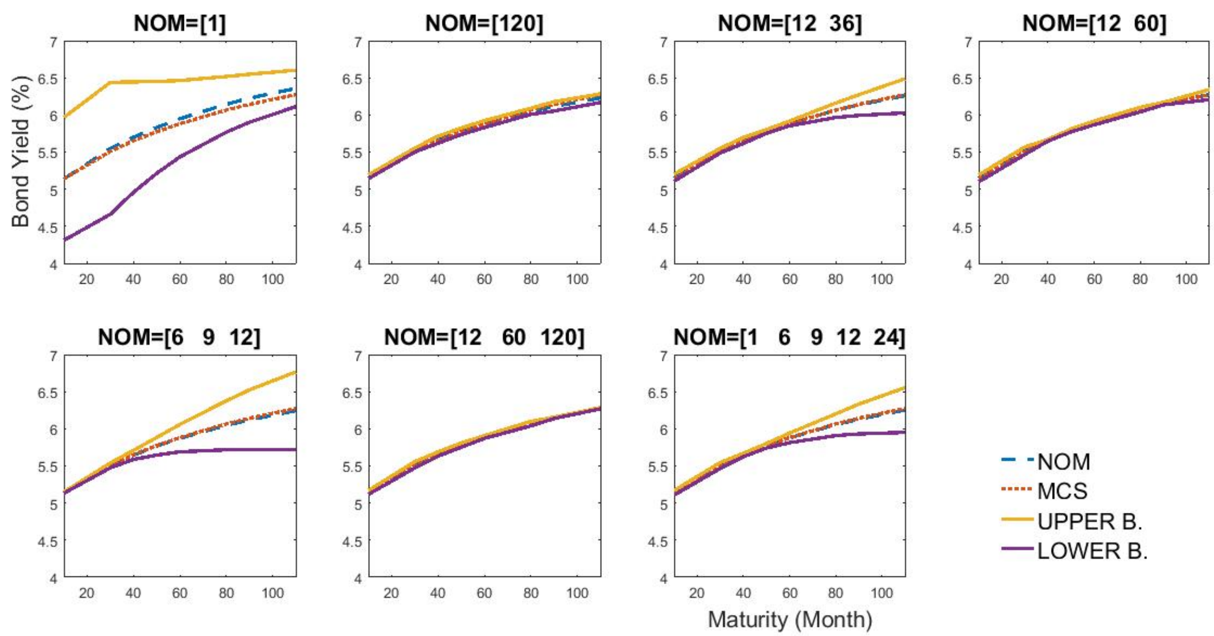


Figure 2.C.3: Yield curves related to different X_{NOM}s, based on $N_0 = 150$ times simulations with the size of $N = 2,000,000$ and $T = 1$. The bound curves of each X_{NOM} are calculated based on the estimated θ^*. The dash lines (legend NOM) represent the expected yields averaged over time, given by the nominal models. The dotted lines (legend MCS) are by the most deviating model in $\mathcal{M}^*_{95\%}$. The solid lines plot the upper bounds and lower bounds.

3

Evaluating the Impacts of Bond Pricing Misspecification on Forecasted Funding Ratio

3.1 Introduction

This paper empirically investigates the impact of misspecification uncertainty on a defined benefit pension funding ratio forecasted 10 years ahead from 2016, using a prediction interval incorporating model misspecification uncertainty.

Typically, a pension fund in the second pillar, supported by both employers and employees, can be classified into three types: a defined contribution type which requires a defined amount of contribution before retirement without promising a defined amount to pay after retirement; a defined benefit type which promises to pay a defined pension every period after retirement with the contribution before retirement specified in an actuarially fair way; and a hybrid type which combines the above two types. This paper looks into the defined benefit type. We consider a set-up in which the defined pensions to be paid form the liabilities of the fund, and the initial assets are set such that their value is equal to the actuarially fair value of the initial liabilities. It is of crucial importance for the pension fund management to monitor closely the funding ratio, i.e., the ratio of assets over liabilities. It helps to perceive the changes of this ratio in the future, so as to adjust policies for preventing the impact from unfavorable outcomes.

Before the financial crisis in 2008, the Dutch pension funding ratio was about 1.40 on average.[1] This average dropped at the crisis, and has been fluctuating between 0.90 and 1.10 since then. The average funding ratio in the third quarter of 2016 is right about 1.00, and in the fourth quarter it is slightly higher. In other words, the pension funding ratio has not yet returned to pre-crisis levels.

A balanced funding ratio shall be maintained being at least one, but preferably more. In other words, assets should at least redeem liabilities. Along the time horizon, the investment return rate affects the assets, and the discount rate is used to value the liabilities. If these two rates are balanced, there is little worry about the funding ratio. Ideally, this could be realized by investing in only fixed-income securities (e.g., bonds) with duration matching the liabilities. In this case, the bond yields on the asset side correspond to the discount effects on the liability side. Because the same rates are used on both sides, the assets and the liabilities are balanced, and the funding ratio remains constant in real terms. However, most pension funds have the ambition to offset the potential loss caused by inflation. To achieve this goal, the fund might want to include stocks as a part of the asset portfolio in order to improve the investment return. As a consequence, the asset side is affected by both bond yields and stock returns. While the liability side is affected only by bond yields, the balance between assets and liabilities is no longer guaranteed. In such a case, in order to forecast the funding ratio, we do need to model stock returns and bond yields.

[1] Source: the Dutch central bank, DNB.

The needs to model stocks and bonds and make forecast raise worries about model uncertainty. We classify model uncertainty into two types. One is parameter uncertainty, and the other is misspecification uncertainty. The former type can be eliminated by increasing the data sample size, if the nominal model would be correctly specified. The estimated parameters are usually evaluated by a classical confidence interval. And a typical prediction interval for the funding ratio forecast can be constructed to indicate where the true outcome is expected to fall with a given significance level. The latter type, however, results from model misspecification. It refers to the situation when the nominal model is essentially misspecified from the true data generating process (DGP).[2] In this situation, simply increasing the sample size cannot eliminate the estimation discrepancy between the nominal model and the true DGP. Put differently, when a nominal model is statistically indistinguishable from the true DGP, only parameter uncertainty exists given a finite sample. But if the nominal model is too far from the true DGP, reasonably misspecification uncertainty appears as well, and considering only parameter uncertainty is no longer sufficient. A good forecast is expected to take into account both types of model uncertainty.

In this paper, we choose to ignore parameter uncertainty, because parameter uncertainty can be eliminated eventually by enlarging the sample size. In the big data era, taking into account model misspecification becomes comparatively more imperative, because the impact of parameter uncertainty could be mitigated with more data information, while model misspecification uncertainty could not. Moreover, in practice, simpler and easier-to-implement models are preferred, even though they might fit the data poorly. This implies that misspecification uncertainty is inevitable in many circumstances. And this is what concerns us the most.[3]

To sum up, when modelling the returns of stocks and bonds, the model un-

[2] In terms of sample, one could consider situation that the nominal model is statistically distinguishable from a pseudo true DGP. "Statistically distinguishable" is the opposite of "statistically indistinguishable" which is used to describe models performing equivalently well based on some statistical criteria, e.g., a test constructed on their log-likelihood functions for a finite sample.

[3] In small samples, parameter uncertainty might play a role. In such a case, it can be incorporated straightforwardly, although finding the (approximate) sampling distribution might be somehow complicated.

certainty inevitably affects the assets, the liabilities, and the funding ratio. To deal with this issue, we incorporates model uncertainty in forecasting the funding ratio. In particular, we focus on the impact of misspecification uncertainty on forecasting bonds, because bonds are the fundamental ingredients affecting both the assets and liabilities, and eventually the funding ratio. The impact of model uncertainty on the funding ratio is considered as transmitted from that on bonds only. We do not consider the model uncertainty of any other elements that could be modelled, such as change in population, survival probabilities, stock returns, etc.

The key to construct a prediction interval incorporating misspecification uncertainty is a misspecification interval. The underlying idea of a misspecification interval is that the expectation, or the forecast, might be derived from models alternative to the nominal one, and it indicates the range of such alternative expectations. In practice, the size of the misspecification interval depends on the discrepancy between the nominal model and the so-call pseudo true DGP, a selected model performing the best empirically. If the nominal model would be far from the true one, the expectation range would be large. Therefore, incorporating a misspecification interval to form a prediction interval makes it more powerful to tell where the true outcome might fall.

In this paper, we consider a class of affine term structure models (ATSMs) to price bonds. This includes three categories: the canonical three-factor model proposed by Litterman and Scheinkman [52], without explicitly imposing the arbitrage-free condition and modelling the evolution of the factors (e.g., VAR(1)); the dynamic Nelson-Siegel model (DNS) by Diebold et al. [16], with the modelling of the factors but not the arbitrage-free condition; and the 3-step OLS model by Adrian et al. [1], with both the arbitrage-free condition and modelling the evolution of the factors. For each category, we postulate some specific models different in terms of estimation approaches and factors. We take the canonical three-factor model as the benchmark nominal model, because it is simple, popular in practice, and is believed to fit the empirical data fairly well. Moreover, the simple structure also implies strong impact of model uncertainty. This could provide us with a clearer effect to analyze the impact.

The research is carried out based on monthly data of the recent 10 years, i.e., from Jan 2007 to Dec 2016. Using the historical data, all postulated models are estimated, and used to select the pseudo true DGPs. We assume that the pseudo true DGPs remain unchanged in the 10 years ahead forecast. We forecast by simulations using Monte-Carlo approach with the Kalman filter.

We set up the evolution of assets and liabilities, for the future 10 years from 2017 to 2026 and the cohort that is 65 years old in 2017 that no longer contributes and has just received its first pension payments. The liabilities are considered up to 30 years from 2017 for the surviving individuals. The survival probabilities estimates are given by the Dutch Royal Actuarial Society, and assumed unaffected by model uncertainty or future changes. Moreover, we impose indexation for inflation: the annual pensions increase yearly by the (assumed to be) constant inflation rate. The value of the initial assets is set equal to the value of the initial liabilities. The assets consist of both stocks and bonds with weights 45% and 55%, respectively. The invested bonds are zero-coupon bonds with 10-year maturity. A buy and sell operation is implemented every period to maintain the 10-year duration. For every period after paying the due pensions, all assets are reinvested. The future stock returns are simulated with mean and variance based on the historical data, while the bond yields are simulated by models. We assume no correlation between bonds and stocks. Additionally, we apply the lower bound restriction on bond yields to be -2%; any bond yield lower than -2% are set to -2%.[4] The funding ratio for each period is formed by the corresponding assets and liabilities.

We need the bond yields with maturities up to 30 years for discounting the liabilities, because the pensions are to be paid for 30 years from 2016. We specifically model the bonds with the maturities of 1-year, 5-year, 9-year, 10-year, 15-year, 19-year, 20-year, 25-year and 30-year. The bond yields with other maturities can be recovered by interpolating the yield curve formed by the modelled bond yields.

The main contribution of this paper is applying a prediction interval incorporating misspecification uncertainty to analyze the forecasts of the funding ratio. We

[4]The -2% lower bound restriction on bond yields corresponds to the nowadays' negative interest rate that is not lower than -2%.

find that this prediction interval is gradually enlarged over the forecast horizon and significantly so, as compared to the prediction interval for the nominal model. In 2026, the forecast of the funding ratio derived from the nominal model is 0.95; its 95% prediction interval predicts that the ratio could fall between 0.66 to 1.36. Yet, the one with misspecification uncertainty shows that it could be between 0.48 and 1.88. The difference between the two types of intervals is considerable and hard to ignore. In such a case, policy makers, such as pension fund investors and regulators, face a much larger decision interval, which suggests more caution and preparations because the funding ratio would fluctuate in a wides range.

In the sensitivity analysis, we choose another nominal model for comparison, namely, the DNS model with yield-only factors estimated by a 2-step approach. This model shows narrower intervals due to a better forecast with the extra assumption on the factors. The forecast by the nominal model is raised to 1.07, meaning that the balance is marginally maintained. The corresponding prediction interval is between 0.85 and 1.33, while the one with misspecification uncertainty is 0.81 and 1.38. The misspecification interval is substantially lessened.

Moreover, we also consider the effect of the duration of the bond portfolio. By including 1-year bonds, we reduce the duration to about 5-year, and by including 20-year bond, we increase the duration to about 15-year. In both cases, the differences between intervals without and with misspecification uncertainty are wider, as compared to the case with 10-year duration. But they are not as significant as in case of changing the benchmark nominal model to the alternative one.

The early works on model uncertainty is reviewed in the book Hansen and Sargent [36]. They combine the ambiguity aversion concept and the max-min preferences in decision theory, and robust control theory in engineering, to study macroeconomics and finance. Decisions are made to maximize the worst case of certain alternative models. The most important approach in their framework is to construct the uncertainty set by the so-called detection error probability approach [2]. This approach, however, by construction only includes statistically indistinguishable models. From our point of view, it addresses only parameter uncertainty, which is not sufficient to explain all uncertainties when there exists a significant discrepancy between the nominal model and the pseudo true one.

The approach in our paper is closely related to the proposal in Schneider and Schweizer [65]. The only difference lies in the determination of the amount of model uncertainty. They apply the model confidence set (MCS) procedure [42] to a set including a number of models, and select a subset containing indistinguishable models that fit empirically the best. The amount of model uncertainty is given by the largest divergence from the MCS. We, instead, measure the divergence straightforwardly from the best-performing one judged by a loss function. We regard this as sufficient to approximate the pseudo true DGP.

This paper is also related to Glasserman and Xu [27], who provide an explicit routine to construct the upper and lower bound models. They apply their approach to risk management, and focus on the worst case as most studies do, while our paper takes into account both the worst and the best cases to form an interval with misspecification uncertainty. This paper contributes to the limited literature on applying such an interval.

This paper is an extension of Li [51], in which misspecification interval is applied to analyze ATMS empirically. There are a few extensions in this paper. Firstly, the interval under consideration is a prediction interval incorporating model uncertainty, rather than a simple misspecification interval. Secondly, this paper takes into account more types of ATSMs and factors, while in Li [51] only yield-only factor models are considered. Thirdly, this paper considers a more complicated process that the impact of model uncertainty transmits, rather than only a simple-10-year annuity as an example. The transmitted impact of model uncertainty on funding ratio is of practical meanings and policy implications for pension fund management.

In terms of the pension topic, Iyengar and Ma [44] propose a robust optimization approach for pension fund management. Unlike our paper, it looks into the problem of pension contribution of firms, and focuses more on the engineering approach to develop. Other pension-related papers considering model uncertainty include Pelsser [62] and Shen et al. [66], who focus on asset pricing and hedging in an incomplete market setting, respectively. Our paper pays attention to the application of a prediction interval with misspecification uncertainty, and analyzes the dynamic change for the future funding ratio.

In the following, we first discuss in Section 3.2 the model uncertainty in theory, the construction of a prediction interval with misspecification uncertainty. Section 3.3 elaborates on the evolutions of assets and liabilities. Section 3.4 presents the types of ATSMs taken into account, the simulation methods to forecast, and the loss function constructed for the pseudo true DGP selection. Section 3.5 contains the empirical analysis. In this section, we choose a benchmark nominal model and the pseudo true DGPs to construct the prediction intervals. We first analyze the bond yields, followed by an analysis of the assets, the liabilities, and the funding ratios. A sensitivity analysis is carried out for an alternative nominal model and different bond durations. Section 3.6 concludes.

3.2 The Construction of Prediction Interval with Misspecification Uncertainty

A prediction interval with misspecification uncertainty describes the range in which the true outcome might fall, with high confidence. It includes the possibility that a nominal model might be misspecified, and the outcome might be obtained from an alternative model. This section discusses our construction of this prediction interval. We firstly describe the framework to evaluate model uncertainty, and then form such a prediction interval in terms of bond yields. We extend this construction to form the prediction intervals with misspecification uncertainty, for the assets, the liabilities, and the funding ratio.

3.2.1 Model Uncertainty Evaluation Framework

An important step in evaluating model uncertainty is to decide a reasonable amount of model uncertainty, an amount that is sufficient to measure the discrepancy between the nominal model and the true DGP.

Anderson et al. [2] propose to use the detection error probability approach. This approach uses empirical data to quantify the amount of model uncertainty and to form an uncertainty set, assuming that the nominal model is statistically indistinguishable from the true DGP. This assumption restricts the uncertainty set such that it accounts for parameter uncertainty only, which can be eliminated when a sufficiently large sample is available. It neglects misspecification uncertainty that cannot be eliminated only by increasing the sample size.

Relaxing the above assumption, Schneider and Schweizer [65] propose to identify pseudo true DGP(s) first from a set of plausible models, so that we have a better idea about the reasonable amount of model uncertainty. This amount of model uncertainty can be measured by the Kullback-Leibler (KL) divergence between the nominal model and the pseudo true DGP(s), quantifying the discrepancy between them.

Given the information provided by the data, they use the model confidence set (MCS) selection procedure [42] to identify a set of best-performing and statistically indistinguishable models from a large collection of plausible models. The MCS is regarded to represent the true DGP. Specifically, the procedure ranks the plausible models in the large set by their performances according to a loss function. The one with the largest value is considered to be the worst-performing one. Each round in the procedure tests whether it is indistinguishable from the rest, at a specified significance level. If not, it will be removed, and the procedure will be imposed on the set of the remaining models. This is repeated until no further exclusion is needed. When only one model remains, it means that the information is sufficient to distinguish a unique best performing model, and this model is simply the one with the smallest value of the loss function. When there is more than one remaining model, the amount of model uncertainty is determined by the largest divergence from the corresponding model in the set.

The MCS ensures the greatest possibility to capture the true DGP because all indistinguishable models are included. In our paper, we ignore the consideration of this greatest possibility, and assume that the information given by the available data is sufficient to rule out all models but the one with the least loss function value.

The amount of model uncertainty is then measured from this unique model. This model is called the pseudo true DGP.

To sum up, we first form an uncertainty set with the KL divergence representing a proper uncertainty amount between the nominal model and the true DGP. This requires a selection from a large set of plausible models, by ranking them according to a loss function that evaluates their performances. A pseudo true DGP whose loss function value is the least is chosen to determine the KL divergence. Secondly, we form the misspecification interval, bounded by the best and the worst expectations of a quantity of interest (e.g., a bond yield), among all expectations derived from the models in the uncertainty set. The models underlying the bounds of the interval are called the "upper bound" and "lower bound" models, respectively. This constructed misspecification interval is the base to incorporate misspecification uncertainty into prediction interval.

In mathematical terms, suppose $g(X)$ is a quantity of interest being a function of a random variable X, assumed to follow the probability distribution $p_X\left(\Phi_X\right)$ with Φ_X a parameter vector. With misspecification uncertainty, one suspects this assumption of X, and claims that X may be ruled by an alternative distribution. Then we evaluate the impact of misspecification uncertainty of p_X on $g(X)$ by a misspecification interval. Here, p_X is the nominal model, whose parameter(s) in Φ_X are estimated by the nominal modelling structure and approach.

Suspecting p_X not being the true DGP, we consider an uncertainty set $\mathcal{P}_X$, which contains alternative models and tries to capture the true DGP as much as possible. Thus, constructing $\mathcal{P}_X$ needs to quantify the discrepancy between the nominal model and the pseudo true DGP given by

$$\widetilde{p}_X \in \left\{\bar{p}_X \in \{p_{X_1}, p_{X_2}, \cdots, p_{X_k}\} \middle| l(\bar{p}_X) \geq l(p_{X_j}), j = 1, \cdots, k\right\},$$

where $l(p_{X_j})$ is the loss function of a model p_{X_j} with Φ_{X_j}, and it is likewise for $l(\bar{p}_X)$.

The discrepancy is measured by the KL divergence defined by $\kappa_j = \mathbb{E}\left(m_j \log m_j\right)$, where m_j is the Radon-Nikodym (RN) derivative defined by the ratio of the alternative model j and the nominal model, $m_j = p_{X_j}/p_X$. Specifically, for the models with a univariate normally distributed random variable, and $\Phi_X = [\mu_X, \sigma_X^2]'$ with μ_X the mean and σ_X^2 the variance of X, the KL divergence κ_j can be calculated by

$$\kappa_j = \log \frac{\sigma_X}{\sigma_{X_j}} + \frac{\sigma_{X_j}^2 + \left(\mu_{X_j} - \mu_X\right)^2}{2\sigma_X^2} - \frac{1}{2}. \tag{3.1}$$

Let the KL divergence κ^* denote the discrepancy between the nominal model p_X and the pseudo true DGP $\widetilde{p}_X$. The model uncertainty set $\mathcal{P}_X$ is then constructed using κ^*, given by

$$\mathcal{P}_X = \left\{p_{X_j} \middle| p_{X_j} = m_j \cdot p_X, \quad \text{and} \quad \mathbb{E}\left(m_j \log m_j\right) \leq \kappa^*\right\},$$

to include any model j whose KL divergence implied by the corresponding m_j is within the constraint κ^*. $\mathcal{P}_X$ can be equivalently expressed in terms of m_j by

$$\mathcal{P}_m = \left\{m_j \middle| \mathbb{E}\left(m_j \log m_j\right) \leq \kappa^*\right\}. \tag{3.2}$$

The expectation of $g(X)$ under p_X is given by $\mathbb{E}\left[g(X)\right]$. The expectation under an alternative model p_{X_j} is given by $\mathbb{E}\left[m_j \cdot g(X)\right]$. From all the expectations derived from the models in the uncertainty set, we are interested in the best and the worst cases. Subject to (3.2), they are obtained by the optimization problems

$$\sup_{m_j \in \mathcal{P}_m} \mathbb{E}\left[m_j \cdot g(X)\right] \qquad \text{and} \qquad \inf_{m_j \in \mathcal{P}_m} \mathbb{E}\left[m_j \cdot g(X)\right], \tag{3.3}$$

respectively. These two quantities form the upper bound and lower bound of the misspecification interval, providing the range within which the true expectation would fall. The corresponding models are the bound models. According to Glasserman and Xu [27], the optimal solution of m_j is given by

$$m_j^* = \frac{\exp\left[\theta \cdot g(X)\right]}{\mathbb{E}\left\{\exp\left[\theta \cdot g(X)\right]\right\}}, \tag{3.4}$$

where θ is in fact the Lagrangian multiplier constraining the condition (3.2) and indicating the impact of misspecification uncertainty on the nominal model. When $\theta > 0$, m_j^* is used for the upper bound model, and when $\theta < 0$, it is used for the lower bound model. As seen, the parameter θ is the crucial ingredient to determine the misspecification interval. We call θ the uncertainty parameter.

Suppose $g(X) = X$ where X is a univariate random variable, and assume that $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ under the nominal model p_X. Since $p_{m_j^*} = m_j^* \cdot p_X$, substituting this in (3.4) shows that X under a bound model $p_{m_j^*}$ follows a normal distribution as well, with mean $\mu_X + \theta\sigma_X^2$, and variance σ_X^2, i.e.,

$$X \sim \mathcal{N}\left(\mu_X + \theta\sigma_X^2, \sigma_X^2\right). \tag{3.5}$$

Using the information in (3.5) and applying (3.1), the KL divergence between a bound model and the nominal model is given by,

$$\kappa_\theta = \frac{1}{2}\theta^2\sigma_X^2. \tag{3.6}$$

Given a nominal model with a specific variance σ_X^2, κ_θ is a function of θ.

By the construction of the problem (3.3), the optimal solution m_j^* exists at the boundary of the constraint (3.2), i.e., $\kappa^* = \mathbb{E}\left(m_j^* \log m_j^*\right)$. Thus $\kappa_\theta = \kappa^*$. After κ^* is obtained by the known distributions p_X and $\widetilde{p}_X$, we are able to solve for the uncertainty parameter θ. The bound models and expectations can then be recovered.

3.2.2 Prediction Interval and Misspecification Interval

In our application, the fundamental quantity of interest is a τ-maturity bond yield at time t, i.e., $g\left(y_t^{(\tau)}\right) = y_t^{(\tau)}$. The nominal model assumes that $y_t^{(\tau)} \sim \mathcal{N}\left(\mu_t, \sigma_t^2\right)$. In the absence of model uncertainty, the outcome of $y_t^{(\tau)}$ is expected to fall in the prediction interval with α significance level given by

$$PI_t\left(y_t^{(\tau)}\right) = \left[\mu_t - z_{1-\frac{\alpha}{2}}\sigma_t, \quad \mu_t + z_{1-\frac{\alpha}{2}}\sigma_t\right] \tag{3.7}$$

where $z_{1-\frac{\alpha}{2}}$ is the value for the $\left(1-\frac{\alpha}{2}\right)$-quantile of a standard Gaussian distribution. For instance, for a 95% prediction interval, $\alpha = 5\%$, and $z_{1-\frac{\alpha}{2}} = 1.96$.

In the presence of misspecification uncertainty, we also consider the prediction intervals of the bound models, as constructed in the previous subsection, rather than that of the nominal model only.[5] The interval describing where the true outcome falls is firstly expanded by the misspecification interval accounting for misspecification uncertainty. This misspecification interval describes the range of expectations for the models in the uncertainty set. Let $\widetilde{y}_t^{(\tau)}$ denote the bond yield under the bound model.[6] By (3.5), the distribution of a bound model is $\widetilde{y}_t^{(\tau)} \sim \mathcal{N}\left(\mu_t \pm \theta_t\sigma_t^2, \sigma_t^2\right)$, where $\theta_t > 0$ is the uncertainty parameter at time t. The misspecification interval of $y_t^{(\tau)}$ is thus bounded by the expectations of $\widetilde{y}_t^{(\tau)}$ derived from the bound models, i.e.,

$$MI_t\left(y_t^{(\tau)}\right) = \left[\mu_t - \theta_t\sigma_t^2, \quad \mu_t + \theta_t\sigma_t^2\right], \qquad \theta_t > 0 \tag{3.8}$$

The prediction interval taking into account misspecification uncertainty is given by

$$MUPI_t\left(y_t^{(\tau)}\right) = \left[\mu_t - \theta_t\sigma_t^2 - z_{1-\frac{\alpha}{2}}\sigma_t, \quad \mu_t + \theta_t\sigma_t^2 + z_{1-\frac{\alpha}{2}}\sigma_t\right], \quad \theta_t > 0. \tag{3.9}$$

In other words, with misspecification uncertainty, the true outcome would fall between the upper bound of the upper bound model's prediction interval, and the lower bound of the lower bound model's prediction interval, with a significance level based on the relevant intervals' construction.

[5] The construction of these bound models is based on the first moments only (as discussed in the previous subsection). The construction of these bound models, taking into account the second moments, and in case of non-normality, higher order moments, is a topic of future research.

[6] It is worthy to note that the observations of $\widetilde{y}_t^{(\tau)}$ and $y_t^{(\tau)}$ are identical, but their underlying distributions are not. In this paper, a variable with $\widetilde{\cdot}$ is denoted as one that follows an alternative distribution.

3.2.3 The Prediction Intervals of the Funding Ratio, with Misspecification Uncertainty

The impact of misspecification uncertainty on the bond yields is transmitted to influence the assets and the liabilities, and eventually the funding ratio. Without making an assumption about the distributions for these variables, we consider their $(1-\alpha)$ prediction intervals given by the $\frac{\alpha}{2}$ and $1-\frac{\alpha}{2}$ quantiles. The prediction intervals of the assets, the liabilities, and the funding ratio based on the nominal model of $y_t^{(\tau)}$ are

$$\begin{aligned} PI_t\left(A_t\right) &= \left[Q_{\frac{\alpha}{2}}\left(A_t\right), \quad Q_{1-\frac{\alpha}{2}}\left(A_t\right)\right], \\ PI_t\left(L_t\right) &= \left[Q_{\frac{\alpha}{2}}\left(L_t\right), \quad Q_{1-\frac{\alpha}{2}}\left(L_t\right)\right], \\ PI_t\left(FR_t\right) &= \left[Q_{\frac{\alpha}{2}}\left(FR_t\right), \quad Q_{1-\frac{\alpha}{2}}\left(FR_t\right)\right], \end{aligned}$$

where $Q_q(A_t)$, $Q_q(L_t)$ and $Q_q(FR_t)$ are the q-quantile of A_t, L_t, and FR_t, respectively.

When taking into account misspecification uncertainty, the prediction intervals based on the bound models of $y_t^{(\tau)}$ become

$$\begin{aligned} MUPI_t\left(A_t\right) &= \left[Q_{\frac{\alpha}{2}}\left(A_t^l\right), \quad Q_{1-\frac{\alpha}{2}}\left(A_t^u\right)\right], \\ MUPI_t\left(L_t\right) &= \left[Q_{\frac{\alpha}{2}}\left(L_t^l\right), \quad Q_{1-\frac{\alpha}{2}}\left(L_t^u\right)\right], \\ MUPI_t\left(FR_t\right) &= \left[Q_{\frac{\alpha}{2}}\left(FR_t^l\right), \quad Q_{1-\frac{\alpha}{2}}\left(FR_t^u\right)\right], \end{aligned}$$

where the variables with the superscript u are derived from upper bound model of $y_t^{(\tau)}$. Likewise, the superscript l is used to indicate the empirical variables derived from the lower bound model of $y_t^{(\tau)}$.

3.3 Assets and Liabilities

This section describes the evolutions of the assets, the liabilities, and the funding ratio. The time line starts at the end of 2016, denoted by $t=0$. We focus on only the cohort of 65 years-old at $t = 0$. At this point, the cohort's labor salary stops, and the cohort has already received its first pension for the coming year. Thus, this first pension payment is excluded in our research.

The liabilities are considered at the end of each period after the annual pensions for the next year are paid. The liabilities sum up the discounted annual pensions to pay. The annual pensions are compensated for inflation, and paid to surviving individuals only. We ignore the condition for inflation indexation, which usually requires the funding ratio to reach a sufficiently high level so as to compensate for the inflation. We assume that the liabilities for this cohort last up to 30 years from $t = 0$.

The initial value of the assets are set to equal that of the initial liabilities, so that the funding ratio is just balanced at $t = 0$. From $t = 0$, the assets are (re)invested in bonds and stocks. The pension fund benefits only from the assets' returns[7] at $t = 1$, and pays the pension. The value of assets at $t = 1$ is considered after the payment. We follow this process for 10 years.

The funding ratio at time t, FR_t, is given by

$$FR_t = \frac{A_t}{L_t}, \quad t = 0, \cdots, 10,$$

where A_t and L_t are the assets and liabilities at time t. When $t = 0$, it is the end of the year 2016, and $t = 10$ is at the end of the year 2026. At $t = 0$, $A_0 = L_0$, and $FR_0 = 1$.

3.3.1 Liabilities

The liabilities at time t aggregate the pensions to be paid in the future, discounted to t. Suppose each individual alive has received the amount of pension π at $t = 0$, and he is promised to receive the same amount annually in future, in real terms. The probability for him to survive τ years and receive pensions is denoted by ${}_{\tau}p_{65,0}^{(g)}$, where g is for the gender. For instance, at $t = 0$, to receive the pension at $t = 1$, an

[7] Because all pension members are already retired, they do not pay contribution anymore.

individual has to survive at least $\tau = 1$ year, with the probability ${}_1p^{(g)}_{65,0}$. Suppose the population of the cohort is $N^{(g)}_{65,0}$ for the gender g. In real terms, the aggregate pensions to be paid at time t are the aggregate pensions to receive for the individuals surviving $\tau = t$ year(s), given by

$$\Pi_t = \sum_g {}_tp^{(g)}_{65,0} \cdot N^{(g)}_{65,0} \cdot \pi, \qquad t = 1, \cdots, 30.$$

Taking into account a constant inflation rate IF, the pension to pay at $t = 1$ would in fact be $(1 + IF) \cdot \Pi_1$. The same annual amount is promised for future. At $t = 2$, the amount to pay is again compensated for the inflation, actualized as $(1 + IF)^2 \cdot \Pi_2$, and this amount becomes the future promise. So on and so forth, the annual pension promised at time t to pay from next period $t + 1$ is thus $(1 + IF)^t \cdot \Pi_t$. The diagram in Figure 3.3.1 shows the flows of the liabilities for each individual alive.

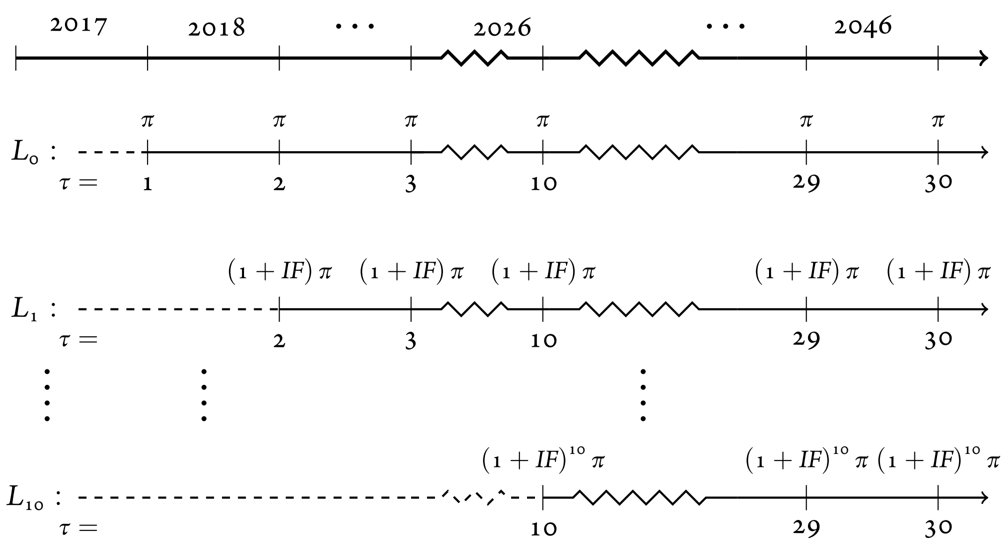


Figure 3.3.1: Liabilities L_t for individual alive at $t = 0, \cdots, 10$

The liabilities L_t are the liabilities after the pensions amount $(1 + IF)^t\,\Pi_t$ is paid, i.e., summing up all Π_τ for $\tau > t$ that are discounted by the τ-term interest rate $y_t^{(\tau)}$

at time t. As a consequence, it is given by

$$\begin{aligned} L_t &= (1+IF)^t\,\Pi_{t+1}\cdot\exp\left(-1\cdot y_t^{(1)}\right)+\cdots+(1+IF)^t\,\Pi_{30}\cdot\exp\left[-(30-t)\cdot y_t^{(30-t)}\right] \\ &= (1+IF)^t\sum_{\tau=t+1}^{30}\Pi_\tau\cdot\exp\left[-(\tau-t)\cdot y_t^{(\tau-t)}\right] \\ &= (1+IF)^t\sum_{\tau=t+1}^{30}\sum_g {}_\tau p_{65,0}^{(g)}\cdot N_{65,0}^{(g)}\cdot\pi\cdot\exp\left[-(\tau-t)\cdot y_t^{(\tau-t)}\right] \\ &= (1+IF)^t\sum_{\tau=t+1}^{30}\sum_g {}_\tau p_{65,0}^{(g)}\cdot N_{65,0}^{(g)}\cdot\pi\cdot P_{b,t}^{(\tau-t)}, \qquad t=0,\cdots 10. \end{aligned}$$

where $P_{b,t}^{(\tau-t)}$ is the price of the bond with time-to-maturity $(\tau-t)$, at time t.

In our research, we consider only the cohort of age 65 in 2016, that has already received its first year pension benefits. The reason for not looking into younger ages is because we do not need to consider the contributions on the asset side, for simplicity. And the reason for not looking into older ages is because we would like mitigate the influence, as the micro-longevity risk[8] on the survival probability ${}_\tau p_{65,t}^{(g)}$ is neglected.

We are also aware that the factor $\left({}_\tau p_{65,0}^{(g)}\cdot P_{b,t}^{(\tau-t)}\right)$ drops dramatically during the 30 years since 65 years-old, and eventually results in a small amount of liabilities afterwards. This is the reason for considering the pension to pay for only 30 years.

3.3.2 Assets

The value of the assets at $t=0$, A_0, by construction satisfies the initial condition $A_0 = L_0$. It is then invested in zero-coupon bonds and stocks. In future periods, after benefiting from the investment with the return rate r_t and paying the pensions $(1+IF)^t\,\Pi_t$, the values of the assets in the next period are given by

$$A_t = A_{t-1}(1+r_t)-(1+IF)^t\,\Pi_t, \qquad t=1,\cdots,10.$$

[8]It refers to non-systematic causes, resulting in, for instance, a smaller and smaller population $N_{65+t,t}^{(g)}$ along the forecast horizon. See Hari et al. [43].

The portfolio return rate r_t is constructed as

$$r_t = w_s \cdot r_{s,t} + w_b \cdot r_{b,t}, \qquad t = 1, \cdots, 10,$$

where $r_{s,t}$ and $r_{b,t}$ are the returns of stocks and bonds, respectively, and w_s and w_b are the weights of the assets allocated to stocks and bonds, respectively. Additionally, $w_s + w_b = 1$ by assumption. The asset investment process is illustrated in the diagram in Figure 3.3.2.

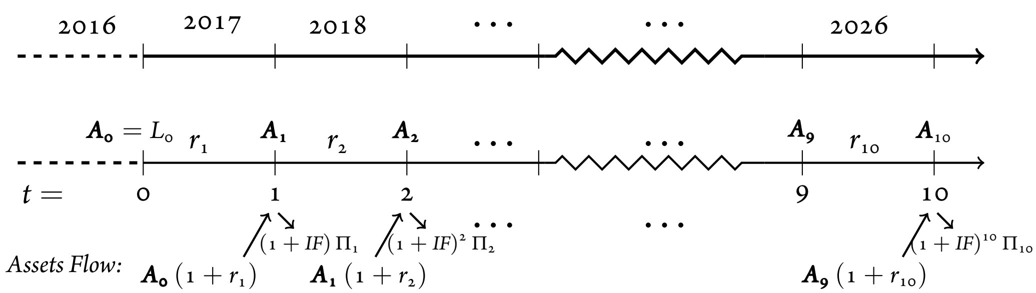


Figure 3.3.2: Pension Assets A_t at $t = 0, \cdots, 10$

In the Netherlands, the investment typically allocates $w_s = 45\%$ to stocks and $w_b = 55\%$ to bonds. The stock gross return $r_{s,t}$ is given by

$$r_{s,t} = \frac{I_{s,t} - I_{s,t-1}}{I_{s,t-1}}, \qquad t = 1, \cdots, 10,$$

where $I_{s,t}$ is the total return index including both the price return and the dividend return.

The investment in bonds consists of only the zero-coupon bonds with 10-year maturity at time t. It is maintained by selling them at 9-year maturity and buying new 10-year zero-coupon bonds at time $t + 1$. That is to say, the duration of the bond investment is sustained as 10 years. The bonds' holding return $r_{b,t}$ is given by

$$r_{b,t} = \frac{P^{(9)}_{b,t} - P^{(10)}_{b,t-1}}{P^{(10)}_{b,t-1}}, \qquad t = 1, \cdots, 10,$$

where $P_{b,t}^{(9)}$ and $P_{b,t-1}^{(10)}$ are the prices of the zero-coupon bonds with 9-year and 10-year maturity, respectively. The prices are determined by the corresponding bond yields $P_{b,t}^{(\tau)} = \exp\left(-\tau \cdot y_t^{(\tau)}\right)$ for τ-year maturity.[9]

3.4 Affine Pricing Models

As discussed, bond yields are fundamental and responsible for the impact of misspecification uncertainty on the funding ratio. They can be modeled in many ways. A significant one is by the class of affine term structure models (ATSMs). Models in this class establish an affine relation between bond yields and factors. These factors could include bond yields with different maturities, macroeconomic variables, latent variables, unspanned variables, etc. Moreover, the ATSMs could be imposed with an arbitrage condition, or estimated by various approaches (e.g., maximum likelihood approach, Kalman filter approach, minimum chi-squared estimation [31], etc.), providing different estimates of the parameters and distributions of bond yields. To study the impact of misspecification uncertainty on bond yields, we include several types of ATSMs: the canonical three-factor model [52], the dynamic Nelson-Siegel model [15, 16], and the arbitrage-free model estimated by the 3-step OLS approach [1]. In this section, we first describe the estimation of the model. The second subsection describes the performance evaluation to select the pseudo true DGP with historical data. In the end, we discuss the simulation approaches used to forecast for different types of models.

3.4.1 Description of Models

This part describes the setting and the estimation of the types of models under consideration.

---

[9]Alternatively, given the total return index $I_{s,t}$, and the bond yields $y_t^{(\tau)}$, one may obtain the stock gross return by $r_{s,t} = \log \frac{I_{s,t}}{I_{s,t-1}}$, and the bond's holding return by $r_{b,t} = \log \frac{P_{b,t}^{(9)}}{P_{b,t-1}^{(10)}} = 10y_{t-1}^{(10)} - 9y_t^{(9)}$.

The Canonical Three-Factor Model

The canonical three-factor model is proposed by Litterman and Scheinkman [52]. It explains the bond yields using the first three principal components as the factors. Many sequential studies find it fitting the empirical data rather well; see Dai and Singleton [13], Driessen et al. [17], Piazzesi [63, 64], et al. It is popular also for its high tractability in practice. The model is postulated as

$$y_t^{(\tau)} = A_\tau + B'_\tau F_t + \varepsilon_t,$$

where $y_t^{(\tau)}$ is the yield with τ-period maturity, $F_t \in \mathbb{R}^{3\times 1}$ is the vector compiling the three factors, and ε_t is the estimation error with zero mean assumption.

The key point of this model is forming the factors in F_t, the first three principal components, from the sample. The sample data typically includes k bond yields with different maturities, including the one(s) to be priced. It is found in the literature that the first three principal components already account for at least 90% of the variation in the bond yields. Thus, it is suggested to use the first three principal components as the factors F_t, referred to as level, slope, and curvature, respectively. B_τ is a vector of the loadings for the three factors. The parameter A_τ is the intercept.

The Dynamic Nelson Siegel Model

The dynamic Nelson Siegel (DNS) model [15, 16] is an affine model without imposing an arbitrage-free restriction. This model is built on the parsimonious approximation of bond yields by Nelson and Siegel [60]. This approximation is a linear formula associated with some latent parameters, weighted by the quantities related to the maturity τ. The DNS regards the dynamic form of the estimated latent parameters as the factors representing the level L_t, the slope S_t, and the curvature C_t, and regards the constructed τ-specific parameters constructed as the loadings B_τ. In particular, the bond yield is given by

$$y_t^{(\tau)} = B'_\tau F_t + \varepsilon_t, \tag{3.10}$$

where

$$B_\tau = \left(1, \quad \frac{1-e^{-\lambda\tau}}{\lambda\tau}, \quad \frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)', \qquad F_t = \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix}.$$

Moreover, Diebold et al. [16] assume that the factors follow a VAR(1) process. For convenience, the VAR(1) process is formulated by mean-centered factors, with the factor means μ_L, μ_S and μ_C, respectively, given by

$$(F_t - \mu) = \Psi\,(F_{t-1} - \mu) + \eta_t, \tag{3.11}$$

where $\mu = \left(\mu_L, \mu_S, \mu_C\right)'$, and the parameter Ψ is estimated by the dynamic factors. The disturbances ε_t and η_t are white noises, orthogonal to each other and to the initial states,[10]

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim WN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_\varepsilon & 0 \\ 0 & \Sigma_\eta \end{pmatrix}\right], \qquad \mathbb{E}(F_0\eta_t') = 0, \qquad \mathbb{E}(F_0\varepsilon_t') = 0.$$

In particular, the Σ_ε matrix is assumed to be diagonal, implying that zero correlation between the bond yields of different maturities. The Σ_η matrix is assumed non-diagonal, allowing correlation between the factors. λ is a fixed parameter chosen such that the loading for C_t is maximized in the approach by Nelson and Siegel [60].

The factors in this model may include yields only, as well as with extra macro-economic variables. In case of yield-only factors, the latent factor F_t is estimated by multiple linear regressions for a number of dependent variables $y_t^{(\tau)}$, with the designed B_τ in (3.10). When k macro-economic factors are included, the new factor vector $\widetilde{F}_t \in \mathbb{R}^{(3+k)\times 1}$ is complied with the estimated F_t and the macro-economic variables. The new loadings matrix $\widetilde{B}_\tau$ is an augmented B_τ by appropriate rows of zeros. The estimator of the parameter Ψ for $\widetilde{F}_t$ is constructed the same way as for

[10]In case that ε_t and η_t are one-dimensioned, we replace Σ_ε^2 and Σ_η^2 by σ_ε^2 and σ_η^2.

F_t in (3.11). In other words, the macro-economic factors do not affect the bond yield pricing process (3.10), but only the VAR(1) process (3.11). Such factors are called unspanned factors; see Duffee [18], Joslin et al. [46], Wright [71], etc.

This model can be estimated by straightforwardly applying OLS to (3.10) and (3.11). In this approach λ is typically set to 0.0609, as in Diebold and Li [15]. Alternatively, one may employ a one-step Kalman filter approach.When long-term liabilities greatly increase in a low interest rate, and the probability of very low yields should be modeled accurately, the Kalman filter approach is preferred.

The Arbitrage-Free ATSM

An arbitrage-free ATSM imposes the arbitrage-free restriction, besides modelling the evolution of the factors. It can be estimated by many approaches. Here we apply the 3-step OLS approach developed by Adrian et al. [1].

The factors are assumed to follow a VAR(1) process,

$$F_t = \mu + \Psi F_{t-1} + \eta_t, \tag{3.12}$$

where μ and Ψ are the parameters characterizing the autoregressive factor vector F_t. The arbitrage-free condition rules out arbitrage opportunities, and guarantees a unique price in a complete market. This price can be used to replicate other contingent claims in the market, a property favored in the financial theory. The replication is implemented through the pricing kernel, which discounts a bond's future price such that the expected discounted price equals today's bond price. In this model, the pricing kernel is assumed to be exponentially affine to the price of risk, and the price of risk is affine to the factors. In this way, the bond yields can be expressed by the recursively determined A_τ and B_τ,

$$y_t^{(\tau)} = A_\tau + B_\tau' F_t + \varepsilon_t, \tag{3.13}$$

where $\varepsilon_t \sim \mathcal{N}(0, \Sigma_\varepsilon)$. All parameters, including the ones in the VAR(1) process, the pricing kernel and the price of risk, are estimated by the 3-step OLS approach.

We refer to Adrian et al. [1] for more details to the set-up and estimation.

The number of factors is not restricted. One may include any yield factors conjectured to affect the bond being priced. One can also include unspanned macroeconomic factors that affect only the VAR(1) process (3.12) but not the affine pricing process (3.13), by adjusting the model structure and estimation as appropriate. In this paper, we take into account a number of such models containing a variety of factors.

3.4.2 Evaluation of the Models

We evaluate a large set of plausible models that includes the three types as described, in order to select the best-performing one to be the pseudo true DGP. Using historical data, a model's performance is judged according to a loss function that compares the historical realizations and estimates.

A typical loss function is built on the likelihood function. In this paper, for a sample Z collecting T-length time series of a bond yield with τ-maturity $\left(y_1^{(\tau)}, \cdots, y_T^{(\tau)}\right)$, the conditional quasi-log-likelihood function is

$$\log l(\varphi) = \frac{T}{2} \log \sigma_\varepsilon^2 - \frac{1}{2} \sum_{t=1}^{T} \frac{\left(y_t^{(\tau)} - A_\tau - B_\tau' F_t\right)^2}{\sigma_\varepsilon^2},$$

where $\varphi = \{A_\tau, B_\tau, \sigma_\varepsilon^2\}$. The loss function is constructed by

$$Q(Z, \varphi) = -2 \log l(\varphi) = T \log \sigma_\varepsilon^2 - \sum_{t=1}^{T} \frac{\left(y_t^{(\tau)} - A_\tau - B_\tau' F_t\right)^2}{\sigma_\varepsilon^2}.$$

A smaller $Q(Z, \varphi)$ indicates better performance, and closer to the true DGP. We rank the values of the loss function given by all the models considered. The model with the least loss function value is regarded as the best-performing one, and used as the psuedo true DGP.

3.4.3 Simulation of the Models

Based on a model, the prediction of bond yields is implemented by simulation. Given the types of model under consideration, we apply two simulation methods, depending on the imposed conditions.

The first simulation method is historical simulation. It is applied to simulating the canonical three-factor model. This model imposes no assumption on the evolution of the factors, and no arbitrage-free condition. In other words, we know little about the distribution of the factors. Therefore, one possible simulation method is the historical simulation, i.e., drawing samples with replacement from the historical data.

More precisely, the historical factors are assumed to satisfy

$$F_t = \mu + \eta_t,$$

where μ is the expectation of the historical data F_t, estimated by $\widehat{\mu} = \frac{1}{T}\sum_{t=1}^{T} F_t$, and $\eta_t = F_t - \widehat{\mu}$ is the error following the empirical distribution. In practice, simulating F_t is simply drawing with replacement from the historical record. Suppose the simulated factors are $\widetilde{F}_t$. Then the simulated bond yields $\widetilde{y}_t^{(\tau)}$ can be calculated with the estimated characterizing parameters $\widehat{A}_\tau$ and $\widehat{B}_\tau$, i.e., $\widetilde{y}_t^{(\tau)} = \widehat{A}_\tau + \widehat{B}_\tau' \widetilde{F}_t$. Since the simulate is based on a random draw, no initial values are assumed.

For other types of models, the factors' behavior is assumed. Therefore, we can forecast by the Monte-Carlo approach. We use the Kalman filter in the simulation. Using the estimated characterizing parameters pertaining to the affine bond yield and VAR(1) factors processes, this approach allows us to simulate bond yields and factors together in one step. The initial value to simulate bond yield is the end-period value in the sample. Simulating N paths, the forecast is simply the mean of the simulated paths.

3.5 Empirical Analysis

Considering that we need the information of bond yields with maturities up to 30 years for the liabilities' analysis, and particularly the maturities of 9 years and 10 years for the assets' analysis, we choose to model the bond yields with maturities of 1-year, 5-year, 9-year, 10-year, 15-year, 19-year 20-year, 25-year, and 30-year. The bond yields with other maturities are recovered by interpolating the yield curve formed by the modelled bond yields.

This section first describes the basis data information, followed by determining the pseudo true DGPs and the prediction intervals of bond yields when misspecification uncertainty is incorporated, and comparing with the prediction intervals without such uncertainty. The next part analyzes the impact of misspecification uncertainty transmitted to the assets, the liabilities, and the funding ratio. In the last part, we conduct a sensitivity analysis, using an alternative nominal model and adjusting the duration of the bond investment.

3.5.1 Data Description

Since we model bond yields, the data needed are the data for factors, including bond yields and macroeconomic variables. Stock returns are needed as well, to predict the investment returns. We collect the recent 10 years' data, i.e., from January 2007 until December 2016. The data are of monthly frequency, with 120 observations in total.

We assume that the yield factors are the European AAA bond yields, whose information is collected in terms of the Nielson-Siegel-Svensson (NSS) parameters from the European Central Bank (ECB). The NSS parameters allow us to generate the annualized AAA bond yield for any maturity. Macroeconomic factors include the Dutch CPI, OECD year over year GDP growth rate, and the ECB fixed rate tenders used for the main refinancing operations (MRO). These variables represent inflation, economic growth and monetary policy, which are macro-economic indexes. We assume that the consumption in retirement occurs in the domestic market only, and is thus affected by the Dutch CPI. We consider OECD year over year GDP growth rates instead of the Dutch GDP, because the bond yields to be modelled are affected by more than just the Dutch economy. The raw data of the fixed rate tenders are documented from Oct 2008; missing observations are replaced by zero in the ECB data. The stock return is obtained by the STOXX 600 total return index.

Table 3.A.1 reports the descriptive statistics for part of the data ready to apply. It does not report all the NSS-estimated bond yields to be used as factors, but only those with the same maturities for the bonds to be modelled. We use 44 bonds with different maturities including the reported ones, to construct the principle components for the factors in the canonical three-factor model. These bonds are with maturities of 1, 3, 6, 9, 10, 15, 18, 21, 30, 40, 50, 80, 90, 110 months, and 1 year to 30 years.

Figure 3.A.1a plots the reported NSS-estimated yields, and Figure 3.A.1b shows the historical data of the stock return and the macroeconomic variables. The NSS-estimated bond yields generally decrease. The 1-year yield fell below zero since 2013, reaching almost -1% in the recent months. The 5-year bond yield also fell below zero since 2015. The plot for the stock return shows fluctuations around a mean value. We assume the stock return normally distributed.

3.5.2 The Prediction Intervals of the Bonds, with Misspecification Uncertainty

To incorporate misspecification uncertainty into the prediction interval of a specific maturity bond yield for the next 10 years ahead, we firstly determine the pseudo true DGP for this bond using the past 10 years' information, and assume the pseudo true DGP remains unchanged for the future 10 years. We determine the amount of model uncertainty and uncertainty parameters in the second part. The uncertainty parameters allow us to incorporate misspecification uncertainty into the prediction interval for this bond. The same implementation is carried out for the other specific maturities bonds.

The Pseudo True DGPs and the Benchmark Nominal Model

The pseudo true DGP for a specific maturity bond is selected from a large set of empirically relevant models, namely the ones described in Section 3.4.1. The factors of the first three principle components in the canonical three-factor model are constructed by the 44 bond yields mentioned earlier. For the DNS model using the yield-only factors, the factors are the reported bond yields. For the DNS model using the macro-yield factors, they are the reported bond yields and macroeconomic variables. For arbitrage-free ATSMs, some factors are only yields by choice, and some includes macroeconomic variables as well. The recent 10-year data are used to estimate the parameters for each empirically relevant model in the selection set. The value of the loss function for each model is obtained by the estimates of these parameters.

Table 2.C.2a collects 45 empirically relevant models, and reports the rankings of their performances by the loss function values in descending order. The pseudo true DGP for each maturity is ranked in the last place. The pseudo true DGPs for all the maturities are assumed unchanged in the 10 years forecast. All the pseudo true DGPs are trending upward in prediction, indicating mean-reverting processes. Figure 3.A.2 illustrates such a process, by including the historical data and the forecast using the DNS models estimated by Kalman filter.

We choose the canonical three-factor model as the benchmark nominal model to model bond yields. The first main reason is that it is popular and easy to implement. It fits empirical data fairly well, and typically explains more than 90% of the variation in the bond yields. The second reason is that the considerable model uncertainty implied by its simple structure could help in studying the impact of model uncertainty. This model imposes no assumption on the evolution of the factors, and no arbitrage-free condition. As a consequence, it might allow for considerable misspecification uncertainty, and cause problematic forecast.

The second reason can be verified by recalling the method to forecast by this model. The employed historical simulation method draws an observation with replacement from the historical data for each period and each simulation path. The forecast of each period is simply the mean of the drawn observations for each period in all paths. As long as the number of paths is large enough, the forecast is constant over time. Hence, there could be a big discrepancy from the mean-reverting pseudo true DGP to this nominal model in forecasting. Consequently, we expect to find large amount of model uncertainty κ_t^*, large values of θ_t and wide prediction intervals with misspecification uncertainty, providing a clearer result to study the impact of model uncertainty.

In other words, the first reason represents practitioners' preference, while the second reason fits our study purpose.

The Prediction Intervals of Bond Yields with Misspecification Uncertainty

We apply monthly data, and predict monthly the bond yields, their prediction intervals without and with misspecification uncertainty respectively, for the future 10 years.

As mentioned, the bond yield for each specific maturity τ is modelled and forecasted individually. The forecast of the bond yield $y_t^{(\tau)}$ is based on $N = 50,000$ simulation paths of 120 months ahead starting from Jan 2017, using the data in Dec 2016 as the initial values. With the variance of the simulations for each period t, the prediction interval for the nominal model $PI_t\left(y_t^{(\tau)}\right)$ can be recovered by (3.7)
.

The prediction interval of $y_t^{(\tau)}$ with misspecification concern, $MUPI_t\left(y_t^{(\tau)}\right)$, combines the misspecification interval and the prediction intervals of the bound model models. It is constructed by the uncertainty parameter θ_t, the forecast μ_t and the forecast variance σ_t^2. The uncertainty parameter θ_t implies the impact of the model uncertainty amount κ_t^* at time t, between the nominal model and the pseudo true DGP for τ-maturity bond. As κ_t^* is obtained by (3.1) with the forecasts and the variances of the nominal model and the pseudo true DGP, θ_t is recovered as well by (3.6), given the condition $\kappa_t^* = \kappa_{\theta_t}$.

Table 3.A.3 reports the estimates of forecasted yield for each year, $\widehat{\mu}_t$, and the

standard deviation $\widehat{\sigma}_t$. The former is generally increasing over maturity, while the latter is decreasing. But both of them remain almost stable over time. Table 4.5.1 reports the estimates $\widehat{\kappa}_t^*$ and $\widehat{\theta}_t$ for the end of each year. Because the pseudo true DGP for each maturity remains the same for the whole forecast horizon, we can compare the size of model uncertainty over time. For instance, for the 1-year maturity bond, the amount of model uncertainty $\widehat{\kappa}_t^*$, decreases dramatically until 2022, and then increases slightly until the end of the horizon. Since $\widehat{\sigma}_t$ is fairly stable over time, correspondingly $\widehat{\theta}_t$ shows the same pattern as $\widehat{\kappa}_t^*$. For the 30-year maturity bond, $\widehat{\kappa}_t^*$ is too small to observe model uncertainty. Thus, the resulting prediction interval with misspecification uncertainty is supposed to be almost identical to that without misspecification uncertainty. $\widehat{\kappa}_t^*$ and $\widehat{\theta}_t$ are remarkably large in the near future and for shorter maturities, but much smaller in the further future and for longer maturities.

The $PI_t\left(y_t^{(\tau)}\right)$ and $MUPI_t\left(y_t^{(\tau)}\right)$, are obtained by (3.7) and (3.9) respectively. Since $\widehat{\theta}_t$ is generally decreasing over time and over maturity, the misspecification interval width determined by $\widehat{\theta}_t\widehat{\sigma}_t^2$ shrinks over time and over maturity. So is the $MUPI_t\left(y_t^{(\tau)}\right)$.

The resulting $PI_t\left(y_t^{(\tau)}\right)$ and $MUPI_t\left(y_t^{(\tau)}\right)$ are plotted in Figure 3.A.3. Figure 3.A.3a and Figure 3.A.3b show the structure of bond yields from the perspective of maturity and time respectively. The dashed lines indicate the $MUPI_t\left(y_t^{(\tau)}\right)$, while the dotted lines indicate the $PI_t\left(y_t^{(\tau)}\right)$. The gap between them is the impact of misspecification uncertainty. We see that the misspecification impact in the near future is stronger, causing the prediction interval with misspecification uncertainty wider than that of the nominal model. As $\widehat{\kappa}^*$ becomes more insignificant over time, $MUPI_t\left(y_t^{(\tau)}\right)$ becomes closer to $PI_t\left(y_t^{(\tau)}\right)$; see the 2026 plot in Figure 3.A.3b. These observations are in line with the implications from combining the results in Table 3.A.3 and Table 4.5.1.

As discussed, the forecast of the nominal model is obtained by historical simulation. Thus, it is a constant, being the mean of historical data. The forecast of the pseudo true DGP, capturing the mean-reverting trend, is time-varying with a low level of initial yield. Hence there will be a big discrepancy at the beginning. As the yield given the pseudo true DGP picks up and recovers the historical mean level in the further future, the discrepancy shrinks, and therefore there is less and less misspecification uncertainty in the longer run.

3.5.3 Funding Ratio Analysis

This section firstly focuses on analyzing the funding ratio using the benchmark nominal model, namely the canonical three-factor model. Then, we implement a sensitivity analysis considering two aspects. One uses a different nominal model, the DNS model with yield-only factors. Another one changes the duration in the bond investment to be about 5 years and 10 years.

The Impact of Misspecification Uncertainty

The simulations of the nominal model and the bound models are used to simulate the assets, the liabilities, and the funding ratio. Their 95% prediction intervals are decided by the 2.5% and 97.5% quantiles.

In simulations, we set a lower bound constraint for the bond yields to be -2%. Whenever the modelled bond yield is lower than -2%, the rate applied to in assets and liabilities is -2%. This constraint mainly works on the pseudo true DGP(s), particularly at the beginning of the forecast horizon on the lower bounds and for shorter maturity bonds. This is because of the low initial values in the simulations. It is not likely to constrain later as the mean-reverting force drives the yields upwards.

Table 3.A.5 reports the outcomes of the forecast, the 95% prediction interval of the nominal model, and the prediction intervals incorporating misspecification uncertainty. Figure 3.A.5a plots these outcomes.

Firstly, the forecasted value of assets decreases constantly. The 95%-quantile prediction interval for the nominal model is wider at the beginning, but becomes narrower at the end of 2017. This means that the variance of the assets under the nominal model becomes smaller. Secondly, $MUPI_t(A_t)$ behaves similarly, but there is another reason resulting in the wider interval in the beginning. The misspecification uncertainty is stronger at the beginning, seen in the 9-year maturity and 10-year maturity in Figure 3.A.3a. Therefore, the return rates based on the bound models are more distant from those based on the nominal model, and result in a wider misspecification interval in the beginning. Thirdly, the gap between $PI_t(A_t)$ and $MUPI_t(A_t)$ is almost the same in the further future. This means that the assets, given by the nominal model and the bound models, move at almost the same rate. This rate is mainly affected by the bond return $r_{b,t}$, associated with the difference between the 9-year and 10-year bond yields. There is a big difference between these two yields due to the significant misspecification uncertainty at the beginning; see the graphs of 9-year and 10-year bond yields in Figure 3.A.3a. Therefore, the rate is not the same. As the misspecification uncertainty becomes insignificant, their difference becomes stable, and thus $r_{b,t}$ given by the nominal model and the bound models, becomes the same. At last, the prediction intervals are positively skewed.

The liabilities at time t involve the bonds with $(30 - t)$ years to maturity. The misspecification intervals of the bond yields become narrower over time. Together with the discounting effect, the impact of misspecification uncertainty on the liabilities becomes almost unobservable. Both $PI_t(L_t)$ and $MUPI_t(L_t)$ gradually become narrower, due to a smaller variance of the liabilities under the nominal model. Like the asset side, the interval is positively skewed.

The forecast of the funding ratio decreases slowly, and ends up at 0.95 in 2026. The funding ratio might not be sufficient to pay the future liabilities. However, the funding ratio is positively skewed, implying a potential upward trend. Additionally, the $PI_t(FR_t)$ in 2026 tells that the funding ratio may be as low as 0.66. Taking into account misspecification uncertainty, the lower bound of $MUPI_t(FR_t)$ drops to 0.51. The gap is as large as 0.15. On the upper bound side, the difference is even larger. The upper bound of $PI_t(FR_t)$ is 1.36, while the one with misspecification uncertainty might become even 1.88. The difference is as substantial as 0.55, implying that the impact of misspecification uncertainty is non-negligible when using the canonical three-factor model as the nominal model.

Sensitivity Analysis

This part implements the sensitivity analysis. The analysis in terms of the canonical three-factor model shows a substantial impact of misspecification uncertainty on the funding ratio, when the investment strategy is simply buying and selling the 10-year and 9-year bonds, respectively. We would like to conduct a sensitivity analysis to see what would change if using a better structured model and different durations.

An Alternative Nominal Model The DNS model imposes an extra assumption on the evolution of the factors. This helps in performing better in-sample and out-of-sample. The forecast in Figure 3.A.1a is made by this DNS model, showing mean-reverting trends, like the pseudo true DGPs. Consequently, we do not expect the same large amount of model uncertainty as in using the benchmark nominal model; see earlier discussion. When the nominal model is replaced by a better structured model, the both prediction intervals, without and with misspecification uncertainty, are expected to shrink.

Figure 3.A.4 plots the forecasts of the bond yields, their prediction intervals of the alternative nominal model and the bound models, using the DNS model with yield-only factors. We see more impact of misspecification uncertainty is decreasing over maturity, but increasing over time, unlike the case using the benchmark nominal model which is decreasing over time only.

Table 3.A.6 reports the forecasts of the assets, the liabilities, and the funding ratio, as well as their prediction intervals with and without misspecification uncertainty. Figure 3.A.5b shows the corresponding graphs. We find that for the assets and liabilities, the prediction intervals with and without model uncertainty are very close. The difference is more obvious in the funding ratio. The difference grows larger over time, but not as dramatically as in using the benchmark model.

The forecasted funding ratio ends with 1.07 at the end of 2026. Compared to those under the benchmark nominal model, both the prediction intervals with and without misspecification uncertainty are narrower, indicating more precise estimation. Like the benchmark model, both the prediction intervals are positively skewed, implying a force of moving upwards. But the median is almost the same as the mean.

However, the impact of the misspecification uncertainty is still significant for the further future. The interval difference is 0.05 for the upper bounds, and 0.04 for the lower bounds. These are not easily neglected amounts, in particular, in case that the pension fund has not completely recovered from the financial crisis.

Investment with Other Durations We also analyze the impact when changing the duration in the bond investment. We investigate two cases, shortening the duration, and lengthening it. In both cases, the liabilities are not affected. The duration change concerns only the asset side, and hence affects the funding ratio.

In the first case, we allocate 30% of the assets to 1-year bonds, and keep 25% to invest in 10-year bonds. The duration becomes approximately 5 years. The 1-year bonds mature the next period, increasing the liquidity of the pension fund. We apply this investment strategy to both the benchmark and alternative nominal models. The results are reported in Table 3.A.7 and Table 3.A.9, and plotted in Figure 3.A.6 and Figure 3.A.8. We see that the intervals become wider for both nominal models. This is mainly because more misspecification uncertainty is embedded in the 1-year bonds; see the graphs in fFigure 3.A.3 and 3.A.4. The differences are not particularly remarkable, although the changes for the benchmark nominal model is relatively larger than those for the alternative nominal model. Other features remain almost the same as in the original strategy.

The second case considers investing 27% of the assets in 20-year bonds. The allocation to stocks remains the same, and the rest are invested in 10-year bonds. The next period, the bonds are sold and replaced by new bonds with the designed maturities. The duration of the bond investment now becomes about 15 years. The results are reported in Table 3.A.8 and Table 3.A.10, and plotted in Figure 3.A.7 and Figure 3.A.9, respectively. They show that the misspecification interval becomes narrower.

This sensitivity analysis shows that longer duration of bond investment provides narrower prediction intervals, due to smaller variance for longer term bonds and insignificant misspecification uncertainty.

3.6 Summary and Conclusions

This paper studies the impact of misspecification uncertainty on predicting the funding ratio of a Dutch defined benefit pension fund 10 years ahead from 2017 to 2026. The ambition of the pension fund management to offset the inflation and thus to invest in stocks breaks the balance of the funding ratio. Bond yields become crucial to affect the funding ratio through both the asset and liability sides. They could be determined by a number of models, but some simple though clearly misspecified ones are typically preferred in practice. This situation motivates the need to take into account model uncertainty. In this research, we ignore the model uncertainty caused by stock return and survival probabilities. Moreover, when modelling bond yields, we only focus on misspecification uncertainty, but ignore parameter uncertainty.

We assume that the investor prefers the canonical three-factor model to predict bond yields; this model does not model the evolution of the factors, or impose the arbitrage free condition, but is simple and preferred. We also apply a lower bound constraint on the bond yields to stay in line with the real financial market. The empirical analysis shows that the impact of misspecification uncertainty on the bond yields is stronger for short term bonds and the near future. The transmitted impact of misspecification uncertainty on assets, liabilities, and funding ratio is significant. The funding ratio over time is below 1. The prediction interval with misspecification uncertainty can be up to 0.67 wider than the one for nominal model.

The sensitivity analysis is carried out to investigate two aspects. One is to use a better structured model as an alternative nominal model, and the other is to alter the duration of the bond investment. The first analysis dramatically reduces the impact of misspecification uncertainty, and results in a funding ratio above 1 in the forecast. The misspecification interval is reduce to 0.09. The second analysis shows that the influence of the bond investment duration is limited. This sensitivity analysis implies that the fundamental reason for a large misspecification interval of the funding ratio is the poor prediction under a nominal bond yield model.

The bound models in this paper are constructed based on the first moment of a bond yield, in case of normality. They can be constructed based on the second moment as well. In case of non-normality, higher moments are also needed to characterize the bound models. These aspects can be the topics of future research.

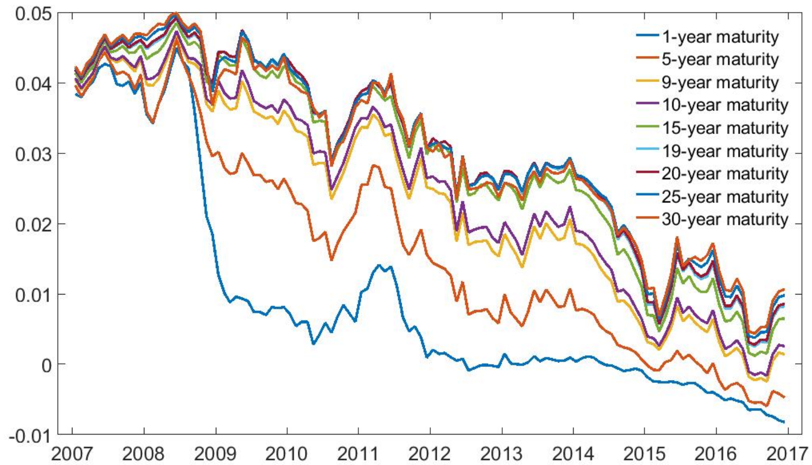
Last but not least, this paper studies the impact of misspecification uncertainty essentially on bonds. The impact on the funding ratio is in fact considered as transmitted from the bond investments in the economy. We do not study the impact on the funding ratio in terms of its distribution, and thus the results are not a straightforward impact of model uncertainty on funding ratio. Moreover, for simplicity, model uncertainty from other economic elements, such as stock returns, demographic change, mortality rates, etc., are not considered. The purpose of this paper is to show the significance of taking into account misspecification uncertainty, and to propose the incorporation of a misspecification interval. These mentioned discrepancies could be improved with extensions, depending on research interests.

Appendix

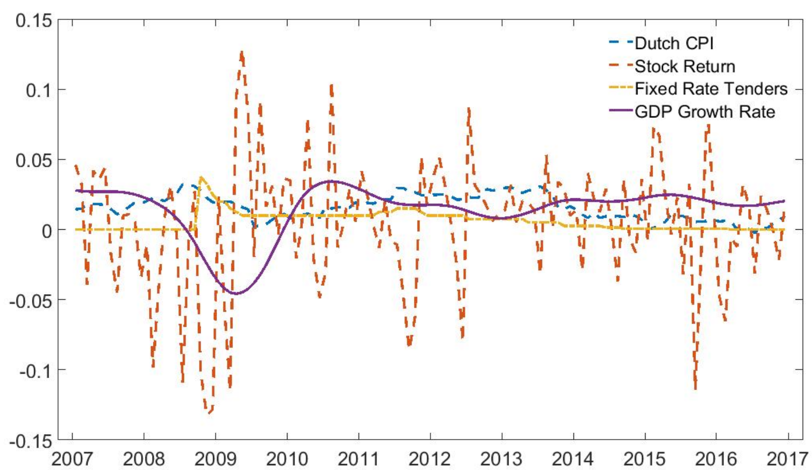
Appendix 3.A Tables and Figures

Table 3.A.1: Data Description. The first panel reports the NSS-estimated yields with τ-year maturity, while the second panel reports the other variables. "STOXX" is short for the stock return, "NLCPI" the dutch CPI, "FRT" the fixed rate tenders, and "GDP Growth" the OECD economic growth rate. The data spans from Jan 2007 to Dec 2016 in monthly frequency.

variables	mean	std	min	max
		NSS Yields with τ-year maturity		
1-year	0.0091	0.0154	-0.0082	0.0449
5-year	0.0167	0.0150	-0.0059	0.0463
9-year	0.0235	0.0144	-0.0024	0.0470
10-year	0.0247	0.0143	-0.0016	0.0473
15-year	0.0287	0.0138	0.0012	0.0485
19-year	0.0301	0.0134	0.0025	0.0491
20-year	0.0303	0.0134	0.0028	0.0492
25-year	0.0306	0.0131	0.0037	0.0496
30-year	0.0303	0.0128	0.0044	0.0499
		Other Variables		
STOXX	0.0039	0.0476	-0.1324	0.1285
NLCPI	0.0158	0.0089	-0.0024	0.0323
FRT	0.0059	0.0067	0.0000	0.0375
GDP Growth	1.3423	1.8529	-4.5800	3.4300



(a) NSS-estimated bond yields



(b) Historical data of stock returns and macro variables

Figure 3.A.1: Descriptive plots of the bond yields, stock returns and other variables. The horizon is from the beginning of 2007 to the beginning of 2017.

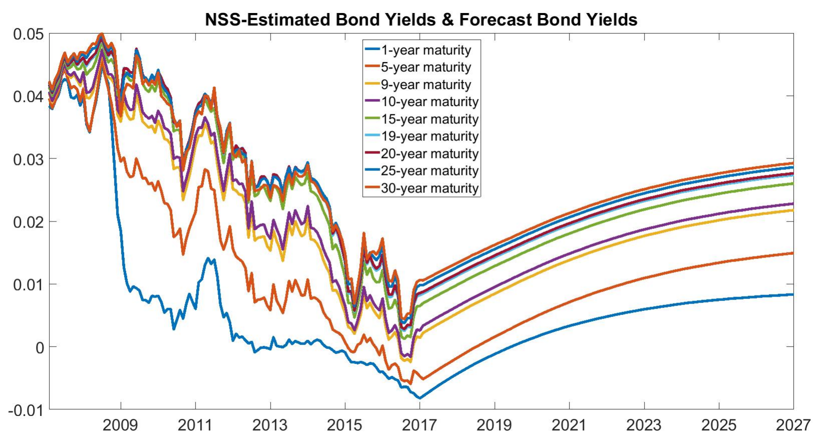


Figure 3.A.2: The NSS-estimated bond yields, and bond yield forecasts by DNS with yield-only factors.

Table 3.A.2: The model rankings according to the values of the loss function.

Index	Model	Macro	τ-year Maturity								
			1	5	9	10	15	19	20	25	30
1	***y3F***	0	42	42	42	42	42	42	42	42	42
2	yYOSSM	0	43	44	44	44	43	39	43	43	43
3	***yYO2S***	0	45	45	43	43	44	43	44	44	44
4	yYMCRSSM	1	44	43	45	45	45	38	45	45	40
5	yYMCR2S	1	25	41	41	41	41	45	39	39	36
6	1/12	1	41	39	39	39	39	44	38	41	41
7	5	1	24	34	35	33	35	40	41	38	45
8	10	1	9	37	33	35	40	35	40	32	38
9	15	1	29	35	34	34	37	37	35	35	35
10	25	1	10	40	37	19	38	32	33	36	32
11	[1/12,5]	1	30	33	40	40	34	33	32	33	34
12	[25,29]	1	32	36	19	38	33	19	34	34	33
13	[10,12.5]	1	8	26	38	37	21	34	37	22	39
14	[1,5,7.5]	1	28	38	23	23	23	25	19	40	37
15	[7.5,12.5,16.7]	1	12	6	25	25	32	24	25	37	22
16	[7.5,20,29]	1	40	12	13	13	19	21	24	21	10
17	[1/12,5,7.5,1/120]	1	27	32	31	24	22	22	21	30	30
18	[5,10,15,20]	1	7	9	36	22	25	26	22	19	25
19	[10,20,25,35]	1	33	29	22	31	15	41	9	25	23
20	[5,7.5,10,20,35]	1	39	30	24	6	24	9	29	10	12
21	[1/12,10,15,20,25]	1	36	10	6	26	18	6	6	29	29
22	[7.5,12.5,15,20,25,35]	1	13	8	26	17	6	30	10	9	9
23	[1/12,5,15,16.7,25,35]	1	19	28	7	21	26	29	26	12	19
24	[1/12,5,7,100,12.5,16.7,20.8]	1	16	31	30	18	36	23	30	6	31
25	[1/12,5,10,15,20,25,35]	1	35	4	27	7	31	36	23	26	28

To be continued

Table 3.A.2 continued

Index	Model	Macro	τ-year Maturity								
			1	5	9	10	15	19	20	25	30
26	1/12	0	15	13	12	27	7	10	31	28	6
27	5	0	26	19	29	36	28	15	28	31	26
28	10	0	31	2	9	30	13	7	7	8	8
29	15	0	38	11	32	11	27	8	8	11	11
30	25	0	37	16	11	12	11	11	11	7	7
31	[1/12,5]	0	18	27	10	29	8	27	27	27	14
32	[25,29]	0	11	7	16	15	12	31	15	23	27
33	[10,12.5]	0	22	3	8	9	10	28	14	14	13
34	[1,5,7.5]	0	6	5	28	32	14	14	36	17	17
35	[7.5,12.5,16.7]	0	21	21	17	10	9	12	12	24	2
36	[7.5,20,29]	0	20	22	15	8	3	3	3	13	4
37	[1/12,5,7.5,1/120]	0	34	15	21	28	5	5	5	2	20
38	[5,10,15,20]	0	14	23	3	16	29	17	17	4	3
39	[10,20,25,35]	0	17	25	5	3	30	13	13	15	5
40	[5,7.5,10,20,35]	0	23	17	1	5	2	16	16	3	15
41	[1/12,10,15,20,25]	0	1	18	20	20	4	1	4	5	1
42	[7.5,12.5,15,20,25,35]	0	3	1	18	1	1	20	1	1	18
43	[1/12,5,15,16.7,25,35]	0	5	24	14	14	17	18	20	18	16
44	[1/12,5,7,100,12.5,16.7,20.8]	0	4	20	4	2	16	4	18	16	24
45	[1/12,5,10,15,20,25,35]	0	**2**	**14**	**2**	**4**	**20**	**2**	**2**	**20**	**21**

Notes: *"y3F"* is short for the canonical three-factor model, *"yYOSSM"* the DNS model estimated by Kalman filter approach with yield factors only, *"yYO2S"* the DNS model estimated by 2-step approach with yield factors only, *"yYMCRSSM"* the DNS model estimated by Kalman filter approach with yield factors and macroeconomic factors, *"yYO2S"* the DNS model estimated by 2-step approach with yield factors and macroeconomic factors. Other models are estimated by 3-step OLS, using only the yield factors with the maturities of the number of *years* indicated in the vectors. The first column provides the index to the model. The *MACRO* column indicates whether macroeconomic factors are included; 1 for yes while 0 for no. The remaining columns show the performance rankings for modelling each maturity bond, in descending order according to the values of the loss function.

t	Year	$\widehat{\mu}_t$ of τ-year maturity $(\times 10^{-2})$								
		1	5	9	10	15	19	20	25	30
0	2016	0.91	1.67	2.35	2.48	2.88	3.01	3.03	3.06	3.04
1	2017	0.91	1.67	2.36	2.48	2.88	3.02	3.03	3.06	3.04
2	2018	0.90	1.66	2.34	2.47	2.87	3.00	3.02	3.05	3.03
3	2019	0.90	1.66	2.34	2.47	2.87	3.00	3.02	3.05	3.03
4	2020	0.89	1.65	2.33	2.45	2.86	2.99	3.01	3.04	3.02
5	2021	0.90	1.67	2.35	2.48	2.88	3.01	3.03	3.06	3.03
6	2022	0.89	1.66	2.34	2.46	2.87	3.00	3.02	3.05	3.02
7	2023	0.92	1.68	2.36	2.48	2.88	3.02	3.03	3.06	3.04
8	2024	0.90	1.66	2.34	2.47	2.87	3.00	3.02	3.05	3.03
9	2025	0.91	1.67	2.35	2.47	2.87	3.01	3.02	3.05	3.03
10	2026	0.90	1.67	2.35	2.47	2.87	3.01	3.03	3.06	3.03

t	Year	$\widehat{\sigma}_t$ of τ-year maturity $(\times 10^{-2})$								
		1	5	9	10	15	19	20	25	30
0	2016	1.54	1.50	1.44	1.43	1.37	1.34	1.33	1.30	1.28
1	2017	1.53	1.49	1.43	1.42	1.37	1.33	1.33	1.30	1.27
2	2018	1.53	1.50	1.44	1.42	1.37	1.34	1.33	1.30	1.28
3	2019	1.53	1.49	1.43	1.42	1.37	1.33	1.33	1.30	1.27
4	2020	1.53	1.49	1.44	1.43	1.37	1.34	1.33	1.30	1.28
5	2021	1.53	1.49	1.43	1.42	1.37	1.33	1.33	1.30	1.27
6	2022	1.52	1.49	1.44	1.42	1.37	1.34	1.33	1.30	1.28
7	2023	1.54	1.50	1.44	1.42	1.37	1.34	1.33	1.30	1.28
8	2024	1.53	1.50	1.44	1.43	1.38	1.34	1.34	1.30	1.28
9	2025	1.54	1.50	1.44	1.43	1.37	1.34	1.33	1.30	1.28
10	2026	1.53	1.49	1.44	1.42	1.37	1.34	1.33	1.30	1.28

Table 3.A.3: This table collects the estimates of forecast $\widehat{\mu}_t$ and standard deviations $\widehat{\sigma}_t$ for τ-maturity bonds, using the benchmark nominal model. The forecasts are obtained by simulation.

t	Year	$\widehat{\kappa}_t^*$ of τ-year maturity $(\times 10^{-2})$								
		1	5	9	10	15	19	20	25	30
0	2016	166.07	154.07	148.71	147.71	138.68	135.51	132.80	126.31	118.22
1	2017	48.58	44.35	41.34	38.37	30.39	34.18	27.90	29.12	35.70
2	2018	21.29	19.07	21.95	14.39	11.63	16.32	10.79	11.41	15.18
3	2019	8.32	9.96	13.54	5.39	4.96	9.38	4.47	4.84	6.37
4	2020	2.36	6.72	8.65	2.36	2.39	5.98	2.18	2.30	3.17
5	2021	0.21	4.93	4.82	1.21	1.04	3.27	1.00	1.07	1.52
6	2022	0.15	3.90	2.23	0.77	0.55	1.51	0.52	0.63	0.79
7	2023	0.68	3.41	0.71	0.44	0.26	0.38	0.24	0.34	0.36
8	2024	1.65	2.87	0.10	0.36	0.21	0.02	0.14	0.22	0.18
9	2025	2.29	2.66	0.02	0.24	0.12	0.12	0.06	0.11	0.06
10	2026	2.96	2.49	0.22	0.20	0.04	0.45	0.03	0.07	0.01

t	Year	$\widehat{\theta}_t$ of τ-year maturity								
		1	5	9	10	15	19	20	25	30
0	2016	118.66	117.29	119.86	120.55	121.32	122.81	122.23	122.02	120.10
1	2017	64.44	63.13	63.41	61.67	57.03	61.96	56.28	58.87	66.32
2	2018	42.56	41.31	46.11	37.69	35.20	42.73	34.93	36.77	43.15
3	2019	26.72	29.96	36.31	23.12	23.05	32.46	22.53	24.00	28.02
4	2020	14.21	24.54	28.91	15.24	15.92	25.79	15.66	16.49	19.67
5	2021	4.27	21.06	21.66	10.93	10.53	19.18	10.68	11.29	13.69
6	2022	3.65	18.72	14.69	8.74	7.66	12.99	7.63	8.59	9.82
7	2023	7.58	17.42	8.27	6.61	5.25	6.55	5.23	6.31	6.65
8	2024	11.85	16.00	3.16	5.98	4.69	1.55	3.95	5.07	4.66
9	2025	13.93	15.38	1.49	4.84	3.52	3.64	2.58	3.60	2.62
10	2026	15.90	14.94	4.66	4.40	2.14	7.11	1.73	2.91	1.28

Table 3.A.4: This table collects the estimated amount of model uncertainty $\widehat{\kappa}_t^*$, and the corresponding uncertainty parameter $\widehat{\theta}_t$, for τ-year maturity bonds, under the benchmark nominal model.

t	Year	Assets				
		$\mu(A_t)$	$Q_{2.5\%}(A_t)$	$Q_{97.5\%}(A_t)$	$Q^l_{2.5\%}(A^l_t)$	$Q^u_{97.5\%}(A^u_t)$
0	2016	30.63	24.83	40.39	20.02	52.24
1	2017	29.27	24.00	36.14	20.56	44.96
2	2018	28.07	22.82	35.35	19.79	43.81
3	2019	26.87	21.52	34.29	18.56	42.76
4	2020	25.64	20.20	33.19	17.14	42.03
5	2021	24.39	18.82	32.10	15.64	41.47
6	2022	23.14	17.53	30.91	14.22	40.61
7	2023	21.88	16.17	29.88	12.82	39.83
8	2024	20.62	14.84	28.67	11.46	38.59
9	2025	19.37	13.50	27.61	10.07	37.68
10	2026	18.08	12.13	26.51	8.62	36.90

t	Year	Liabilities				
		$\mu(L_t)$	$Q_{2.5\%}(L_t)$	$Q_{97.5\%}(L_t)$	$Q^l_{2.5\%}(L^l_t)$	$Q^u_{97.5\%}(L^u_t)$
0	2016	30.63	24.83	40.39	20.02	52.24
1	2017	29.52	23.93	38.54	21.50	44.07
2	2018	28.40	23.15	36.87	21.60	40.15
3	2019	27.30	22.41	35.10	21.36	37.18
4	2020	26.16	21.62	33.34	20.89	34.73
5	2021	24.98	20.69	31.59	20.19	32.50
6	2022	23.80	19.89	29.84	19.52	30.47
7	2023	22.62	18.99	28.11	18.74	28.51
8	2024	21.43	18.07	26.37	17.89	26.63
9	2025	20.24	17.12	24.67	16.96	24.91
10	2026	19.01	16.18	22.98	16.02	23.22

t	Year	Funding Ratio				
		$\mu(FR_t)$	$Q_{2.5\%}(FR_t)$	$Q_{97.5\%}(FR_t)$	$Q^l_{2.5\%}(FR^l_t)$	$Q^u_{97.5\%}(FR^u_t)$
0	2016	1.00	1.00	1.00	1.00	1.00
1	2017	1.00	0.83	1.23	0.82	1.38
2	2018	1.00	0.82	1.23	0.77	1.42
3	2019	0.99	0.81	1.23	0.74	1.47
4	2020	0.99	0.80	1.24	0.70	1.53
5	2021	0.98	0.78	1.25	0.67	1.58
6	2022	0.98	0.77	1.27	0.64	1.64
7	2023	0.97	0.75	1.29	0.60	1.70
8	2024	0.97	0.72	1.31	0.56	1.75
9	2025	0.96	0.70	1.33	0.52	1.81
10	2026	0.95	0.66	1.36	0.48	1.88

Table 3.A.5: This table reports the 95% prediction intervals without and with misspecification uncertainty, for of the assets, the liabilities and the funding ratios. They are built on the bond yields estimated by the **benchmark nominal model**. Results are plotted in Figure 3.A.5a.

t	Year	Assets				
		$\mu(A_t)$	$Q_{2.5\%}(A_t)$	$Q_{97.5\%}(A_t)$	$Q_{2.5\%}(A_t^l)$	$Q_{97.5\%}(A_t^u)$
0	2016	38.18	36.43	40.03	36.37	40.09
1	2017	36.01	32.97	39.23	32.92	39.29
2	2018	33.89	30.27	37.77	30.23	37.82
3	2019	31.90	27.99	36.15	27.98	36.15
4	2020	30.03	25.93	34.51	25.76	34.73
5	2021	28.27	24.02	32.93	23.66	33.41
6	2022	26.60	22.25	31.41	21.78	32.06
7	2023	25.02	20.59	30.00	20.08	30.70
8	2024	23.50	18.96	28.60	18.51	29.22
9	2025	22.04	17.40	27.22	17.04	27.70
10	2026	20.63	15.90	25.94	15.69	26.19

t	Year	Liabilities				
		$\mu(L_t)$	$Q_{2.5\%}(L_t)$	$Q_{97.5\%}(L_t)$	$Q_{2.5\%}(L_t^l)$	$Q_{97.5\%}(L_t^u)$
0	2016	38.18	36.43	40.03	36.37	40.09
1	2017	35.54	31.07	40.72	30.93	40.94
2	2018	32.99	27.95	39.07	27.70	39.46
3	2019	30.69	25.68	36.85	25.34	37.41
4	2020	28.62	23.87	34.49	23.48	35.15
5	2021	26.76	22.32	32.29	21.90	32.99
6	2022	25.05	20.94	30.09	20.49	30.85
7	2023	23.48	19.69	28.13	19.24	28.91
8	2024	22.00	18.57	26.19	18.09	27.00
9	2025	20.59	17.49	24.34	17.04	25.10
10	2026	19.25	16.41	22.58	15.98	23.31

t	Year	Funding Ratio				
		$\mu(FR_t)$	$Q_{2.5\%}(FR_t)$	$Q_{97.5\%}(FR_t)$	$Q_{2.5\%}(FR_t^l)$	$Q_{97.5\%}(FR_t^u)$
0	2016	1.00	1.00	1.00	1.00	1.00
1	2017	1.02	0.93	1.11	0.93	1.11
2	2018	1.03	0.92	1.15	0.91	1.16
3	2019	1.04	0.92	1.18	0.90	1.19
4	2020	1.05	0.91	1.20	0.89	1.23
5	2021	1.06	0.91	1.22	0.88	1.27
6	2022	1.07	0.90	1.24	0.86	1.30
7	2023	1.07	0.89	1.26	0.85	1.33
8	2024	1.07	0.88	1.28	0.84	1.35
9	2025	1.07	0.87	1.30	0.82	1.36
10	2026	1.07	0.85	1.33	0.81	1.38

Table 3.A.6: This table reports the 95% prediction intervals without and with misspecification uncertainty, for of the assets, the liabilities and the funding ratios. They are built on the **alternative nominal model**. Results are plotted in Figure 3.A.5b.

t	Year	Assets				
		$\mu(A_t)$	$Q_{2.5\%}(A_t)$	$Q_{97.5\%}(A_t)$	$Q_{2.5\%}(A_t^l)$	$Q_{97.5\%}(A_t^u)$
0	2016	30.63	24.83	40.39	20.02	52.24
1	2017	29.03	23.79	36.64	19.72	47.06
2	2018	27.55	22.23	35.55	18.45	45.61
3	2019	26.08	20.59	34.36	16.88	44.42
4	2020	24.60	18.90	33.07	15.19	43.33
5	2021	23.10	17.26	31.84	13.48	42.32
6	2022	21.61	15.62	30.58	11.80	41.32
7	2023	20.11	14.05	29.27	10.16	40.16
8	2024	18.62	12.39	27.98	8.50	39.00
9	2025	17.13	10.77	26.70	6.85	37.85
10	2026	15.64	9.17	25.44	5.20	36.76

t	Year	Liabilities				
		$\mu(L_t)$	$Q_{2.5\%}(L_t)$	$Q_{97.5\%}(L_t)$	$Q_{2.5\%}(L_t^l)$	$Q_{97.5\%}(L_t^u)$
0	2016	30.63	24.83	40.39	20.02	52.24
1	2017	29.52	23.93	38.54	21.50	44.07
2	2018	28.40	23.15	36.87	21.60	40.15
3	2019	27.30	22.41	35.10	21.36	37.18
4	2020	26.16	21.62	33.34	20.89	34.73
5	2021	24.98	20.69	31.59	20.19	32.50
6	2022	23.80	19.89	29.84	19.52	30.47
7	2023	22.62	18.99	28.11	18.74	28.51
8	2024	21.43	18.07	26.37	17.89	26.63
9	2025	20.24	17.12	24.67	16.96	24.91
10	2026	19.01	16.18	22.98	16.02	23.22

t	Year	Funding Ratio				
		$\mu(FR_t)$	$Q_{2.5\%}(FR_t)$	$Q_{97.5\%}(FR_t)$	$Q_{2.5\%}(FR_t^l)$	$Q_{97.5\%}(FR_t^u)$
0	2016	1.00	1.00	1.00	1.00	1.00
1	2017	1.00	0.74	1.35	0.70	1.56
2	2018	0.98	0.72	1.35	0.65	1.61
3	2019	0.97	0.70	1.34	0.61	1.65
4	2020	0.95	0.68	1.34	0.57	1.69
5	2021	0.94	0.66	1.34	0.53	1.73
6	2022	0.92	0.63	1.34	0.49	1.78
7	2023	0.90	0.60	1.34	0.44	1.82
8	2024	0.88	0.57	1.35	0.40	1.86
9	2025	0.85	0.53	1.35	0.34	1.90
10	2026	0.83	0.48	1.37	0.28	1.96

Table 3.A.7: This table reports the 95% prediction intervals without and with misspecification uncertainty, for of the assets, the liabilities and the funding ratios. They are built on the **benchmark nominal model** with 30% investment in **1-year bond**. Results are plotted in Figure 3.A.6.

t	Year	Assets				
		$\mu(A_t)$	$Q_{2.5\%}(A_t)$	$Q_{97.5\%}(A_t)$	$Q_{2.5\%}(A_t^l)$	$Q_{97.5\%}(A_t^u)$
0	2016	30.63	24.83	40.39	20.02	52.24
1	2017	29.52	23.93	37.04	21.04	43.81
2	2018	28.51	22.87	36.65	20.47	43.73
3	2019	27.51	21.76	35.81	19.42	42.86
4	2020	26.48	20.64	34.83	18.18	42.24
5	2021	25.41	19.41	33.79	16.87	41.63
6	2022	24.34	18.27	32.91	15.60	41.19
7	2023	23.27	17.13	31.83	14.43	40.21
8	2024	22.19	15.92	30.83	13.21	39.23
9	2025	21.12	14.77	29.88	12.01	38.56
10	2026	20.00	13.51	28.93	10.63	38.24

t	Year	Liabilities				
		$\mu(L_t)$	$Q_{2.5\%}(L_t)$	$Q_{97.5\%}(L_t)$	$Q_{2.5\%}(L_t^l)$	$Q_{97.5\%}(L_t^u)$
0	2016	30.63	24.83	40.39	20.02	52.24
1	2017	29.52	23.93	38.54	21.50	44.07
2	2018	28.40	23.15	36.87	21.60	40.15
3	2019	27.30	22.41	35.10	21.36	37.18
4	2020	26.16	21.62	33.34	20.89	34.73
5	2021	24.98	20.69	31.59	20.19	32.50
6	2022	23.80	19.89	29.84	19.52	30.47
7	2023	22.62	18.99	28.11	18.74	28.51
8	2024	21.43	18.07	26.37	17.89	26.63
9	2025	20.24	17.12	24.67	16.96	24.91
10	2026	19.01	16.18	22.98	16.02	23.22

t	Year	Funding Ratio				
		$\mu(FR_t)$	$Q_{2.5\%}(FR_t)$	$Q_{97.5\%}(FR_t)$	$Q_{2.5\%}(FR_t^l)$	$Q_{97.5\%}(FR_t^u)$
0	2016	1.00	1.00	1.00	1.00	1.00
1	2017	1.00	0.93	1.14	0.93	1.26
2	2018	1.01	0.90	1.15	0.88	1.29
3	2019	1.01	0.89	1.17	0.84	1.35
4	2020	1.01	0.87	1.19	0.80	1.41
5	2021	1.02	0.86	1.22	0.77	1.48
6	2022	1.02	0.84	1.25	0.74	1.55
7	2023	1.03	0.83	1.28	0.71	1.62
8	2024	1.04	0.81	1.32	0.68	1.68
9	2025	1.04	0.79	1.36	0.65	1.75
10	2026	1.05	0.77	1.41	0.61	1.85

Table 3.A.8: This table reports the 95% prediction intervals without and with misspecification uncertainty, for of the assets, the liabilities and the funding ratios. They are built on the **benchmark nominal model** with 27% investment in **20-year bond**. Results are plotted in Figure 3.A.7.

t	Year	Assets				
		$\mu(A_t)$	$Q_{2.5\%}(A_t)$	$Q_{97.5\%}(A_t)$	$Q_{2.5\%}(A_t^l)$	$Q_{97.5\%}(A_t^u)$
0	2016	38.18	36.45	40.02	36.41	40.07
1	2017	36.05	33.61	38.58	33.56	38.64
2	2018	34.01	31.16	37.00	31.14	37.03
3	2019	32.07	28.92	35.40	28.83	35.50
4	2020	30.21	26.82	33.81	26.58	34.11
5	2021	28.44	24.86	32.29	24.47	32.79
6	2022	26.73	22.99	30.81	22.47	31.47
7	2023	25.08	21.17	29.35	20.58	30.13
8	2024	23.47	19.41	27.94	18.77	28.79
9	2025	21.90	17.71	26.56	17.06	27.45
10	2026	20.36	16.04	25.20	15.40	26.08

t	Year	Liabilities				
		$\mu(L_t)$	$Q_{2.5\%}(L_t)$	$Q_{97.5\%}(L_t)$	$Q_{2.5\%}(L_t^l)$	$Q_{97.5\%}(L_t^u)$
0	2016	38.18	36.45	40.02	36.41	40.07
1	2017	35.50	31.04	40.67	30.86	40.95
2	2018	32.97	27.94	39.05	27.65	39.53
3	2019	30.68	25.64	36.81	25.25	37.46
4	2020	28.63	23.89	34.49	23.44	35.23
5	2021	26.77	22.29	32.26	21.84	33.03
6	2022	25.07	20.94	30.14	20.48	30.92
7	2023	23.50	19.70	28.13	19.24	28.91
8	2024	22.02	18.56	26.19	18.10	26.97
9	2025	20.61	17.46	24.35	17.01	25.12
10	2026	19.26	16.46	22.57	16.01	23.31

t	Year	Funding Ratio				
		$\mu(FR_t)$	$Q_{2.5\%}(FR_t)$	$Q_{97.5\%}(FR_t)$	$Q_{2.5\%}(FR_t^l)$	$Q_{97.5\%}(FR_t^u)$
0	2016	1.00	1.00	1.00	1.00	1.00
1	2017	1.02	0.90	1.14	0.90	1.14
2	2018	1.04	0.89	1.20	0.88	1.21
3	2019	1.05	0.88	1.24	0.86	1.26
4	2020	1.06	0.88	1.26	0.85	1.30
5	2021	1.07	0.87	1.29	0.84	1.34
6	2022	1.07	0.87	1.31	0.83	1.37
7	2023	1.07	0.86	1.32	0.81	1.39
8	2024	1.07	0.84	1.34	0.79	1.41
9	2025	1.07	0.83	1.35	0.77	1.43
10	2026	1.06	0.81	1.36	0.75	1.45

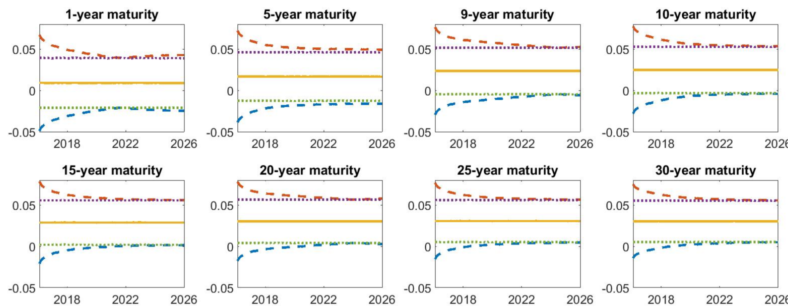
Table 3.A.9: This table reports the 95% prediction intervals without and with misspecification uncertainty, for of the assets, the liabilities and the funding ratios. They are built on the **alternative nominal model** with 30% investment in **1-year bond**. Results are plotted in Figure 3.A.8.

t	Year	Assets				
		$\mu(A_t)$	$Q_{2.5\%}(A_t)$	$Q_{97.5\%}(A_t)$	$Q_{2.5\%}(A_t^l)$	$Q_{97.5\%}(A_t^u)$
0	2016	38.18	36.45	40.02	36.41	40.07
1	2017	35.79	32.00	39.96	31.92	40.06
2	2018	33.47	28.91	38.64	28.86	38.72
3	2019	31.30	26.33	36.87	26.04	37.31
4	2020	29.29	24.20	35.08	23.62	35.94
5	2021	27.40	22.25	33.28	21.44	34.52
6	2022	25.64	20.49	31.60	19.54	33.08
7	2023	23.98	18.80	29.96	17.82	31.50
8	2024	22.40	17.24	28.38	16.29	29.83
9	2025	20.88	15.74	26.86	14.92	28.09
10	2026	19.42	14.27	25.43	13.62	26.32

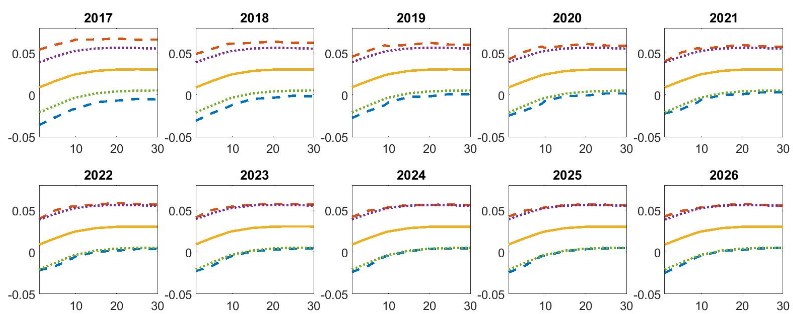
t	Year	Liabilities				
		$\mu(L_t)$	$Q_{2.5\%}(L_t)$	$Q_{97.5\%}(L_t)$	$Q_{2.5\%}(L_t^l)$	$Q_{97.5\%}(L_t^u)$
0	2016	38.18	36.45	40.02	36.41	40.07
1	2017	35.50	31.04	40.67	30.86	40.95
2	2018	32.97	27.94	39.05	27.65	39.53
3	2019	30.68	25.64	36.81	25.25	37.46
4	2020	28.63	23.89	34.49	23.44	35.23
5	2021	26.77	22.29	32.26	21.84	33.03
6	2022	25.07	20.94	30.14	20.48	30.92
7	2023	23.50	19.70	28.13	19.24	28.91
8	2024	22.02	18.56	26.19	18.10	26.97
9	2025	20.61	17.46	24.35	17.01	25.12
10	2026	19.26	16.46	22.57	16.01	23.31

t	Year	Funding Ratio				
		$\mu(FR_t)$	$Q_{2.5\%}(FR_t)$	$Q_{97.5\%}(FR_t)$	$Q_{2.5\%}(FR_t^l)$	$Q_{97.5\%}(FR_t^u)$
0	2016	1.00	1.00	1.00	1.00	1.00
1	2017	1.01	0.95	1.07	0.95	1.08
2	2018	1.02	0.94	1.10	0.93	1.12
3	2019	1.02	0.93	1.12	0.90	1.15
4	2020	1.02	0.92	1.14	0.88	1.19
5	2021	1.02	0.90	1.15	0.85	1.22
6	2022	1.02	0.89	1.17	0.83	1.25
7	2023	1.02	0.87	1.19	0.80	1.28
8	2024	1.02	0.85	1.20	0.78	1.30
9	2025	1.01	0.82	1.22	0.76	1.31
10	2026	1.01	0.80	1.24	0.74	1.32

Table 3.A.10: This table reports the 95% prediction intervals without and with misspecification uncertainty, for of the assets, the liabilities and the funding ratios. They are built on the **alternative nominal model** with 27% investment in **20-year bond**. Results are plotted in Figure 3.A.9.

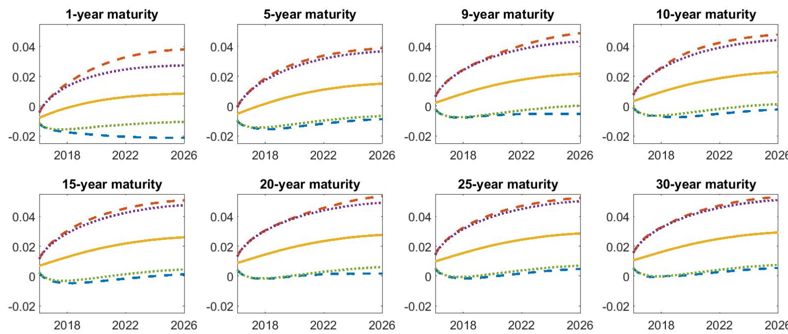


(a) Bond yields over time

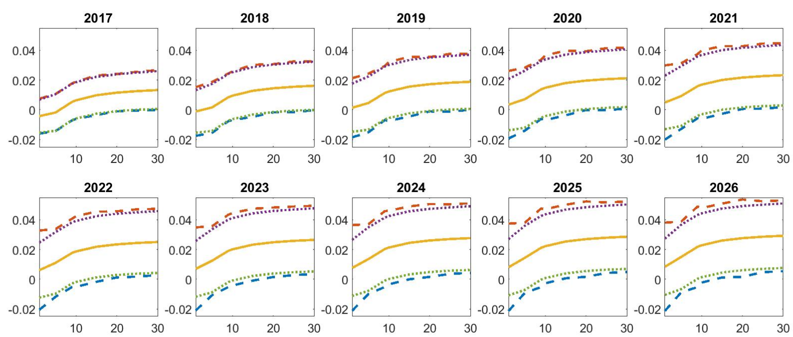


(b) Term structure in the end of the Year

Figure 3.A.3: The prediction intervals with and without model misspecification uncertainty, for the term structure of the bond yields under the **benchmark nominal model**. The dotted lines indicate the 95% prediction intervals for the nominal model, while the dashed lines indicate the prediction intervals with misspecification uncertainty.

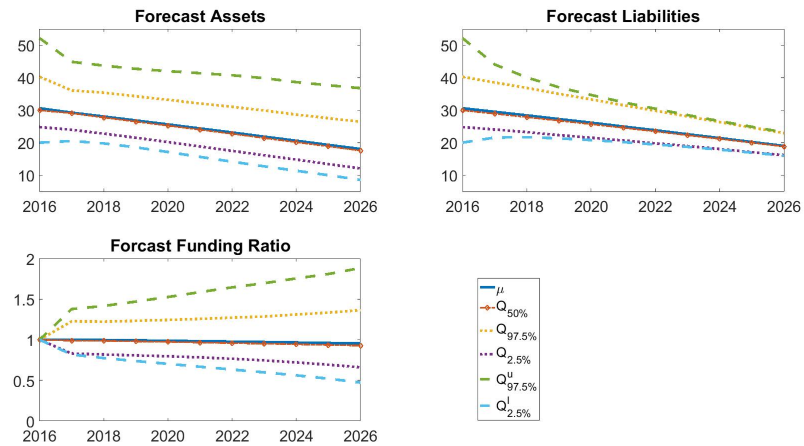


(a) Bond yields over Years

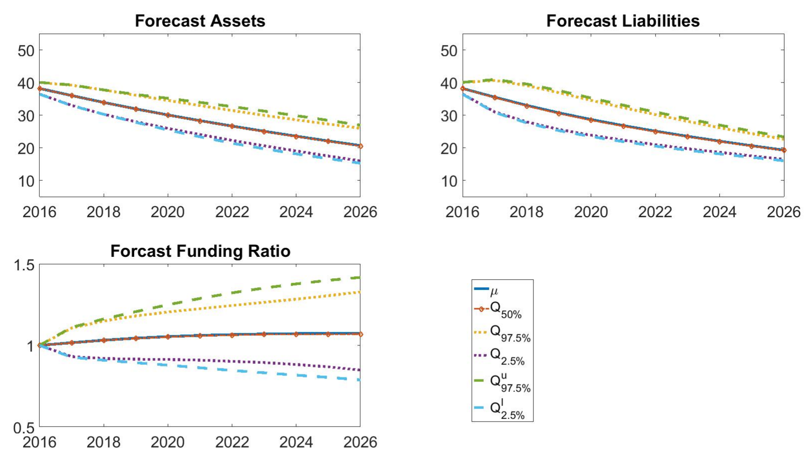


(b) Term structure in the end of the Year

Figure 3.A.4: The The prediction intervals with and without model misspecification uncertainty, for the term structure of the bond yields under the **alternative nominal model**. The dotted lines indicate the 95% prediction intervals for the nominal model, while the dashed lines indicate the prediction intervals with misspecification uncertainty.



(a) Forecasts of the assets, the liabilities and the funding ratio, using the **benchmark nominal model**.



(b) Forecasts of the assets, the liabilities and the funding ratio, using the **alternative nominal model**.

Figure 3.A.5: The results of the forecasts are reported in Table 3.A.5 and 3.A.6, respectively. The solid line and the dashed-diamond line are for the expectations and the medians derived from the nominal model. The dotted lines indicate the 95% prediction intervals for the nominal model, while the dashed lines indicate the prediction intervals with misspecification uncertainty.

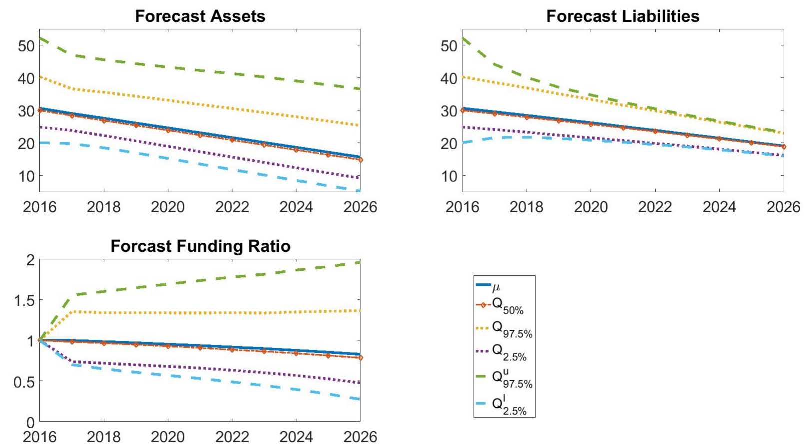


Figure 3.A.6: Forecasts of the assets, the liabilities and the funding ratio, using the **benchmark nominal model**, with 30% **1-year bond** in the investment.

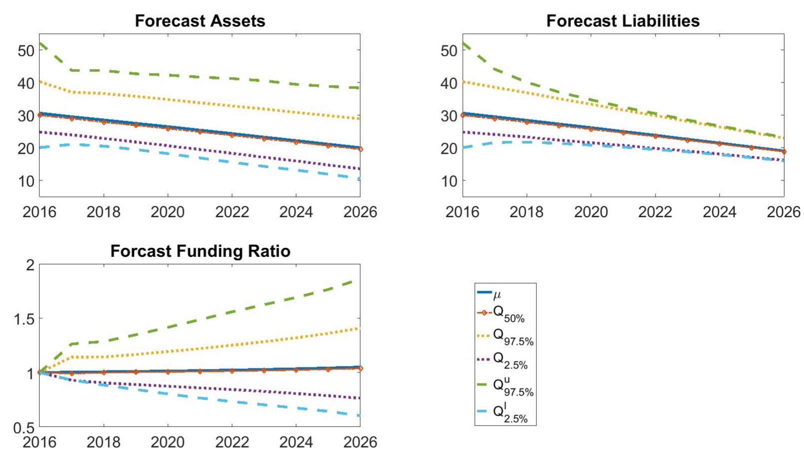


Figure 3.A.7: Forecasts of the assets, the liabilities and the funding ratio, using the **benchmark nominal model**, with 27% **20-year bond** in the investment.

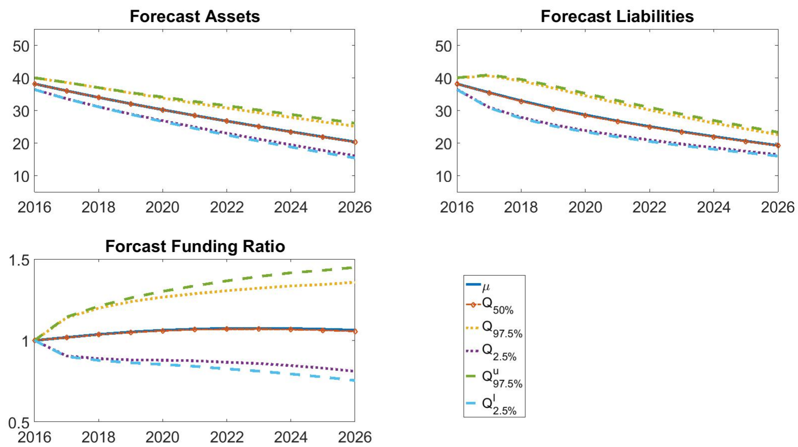


Figure 3.A.8: Forecasts of the assets, the liabilities and the funding ratio, using the **alternative nominal model**, with 30% **1-year bond** in the investment.

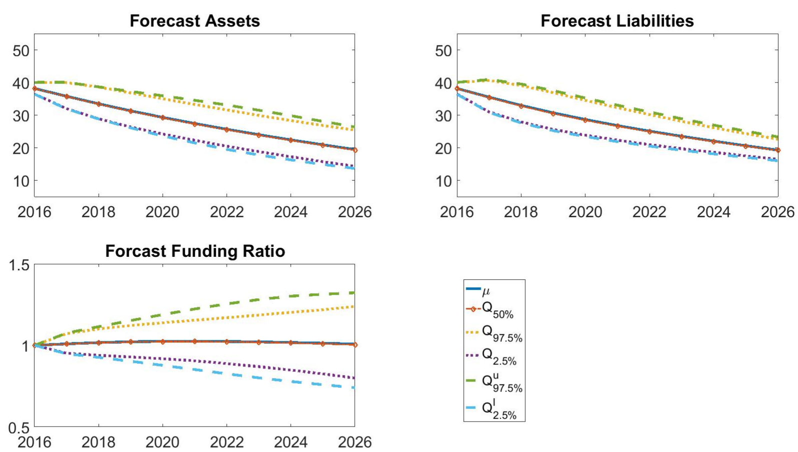


Figure 3.A.9: Forecasts of the assets, the liabilities and the funding ratio, using the **alternative nominal model**, with 27% **20-year bond** in the investment.

4

Portfolio Rules and Interest Rates under Model Uncertainty

4.1 Introduction

This paper studies intertemporal continuous-time investment-consumption decision in a complete market setting, and the asset pricing in equilibrium, where the representative agent worries about model uncertainty, i.e., uncertainty of the probability used for making decisions. Our paper incorporates model uncertainty into a modified Cox-Ingersoll-Ross (CIR) general equilibrium setting Cox et al. [11], and solves it in the spirit of Hansen et al. [41]. This modified setting leads to a decision under commitment. Hansen and Sargent [34] discuss such a problem and propose a time-consistent solution approach. In this paper, we propose a alternative solution approach, and achieve time consistency as well. We provide the closed-form solutions of dynamic investment rules and equilibrium expected assets returns. Moreover, we compare our results with those in the existing literature, in order to understand the impact of model uncertainty under different measures.

We consider a modified Cox-Ingersoll-Ross model (CIR) [11], where the economy is composed of production, consumption, and investment. The market is postulated as being complete. The representative agent decides every period to invest his wealth in a production technology, a risk-free asset, and risky assets. Unlike the original CIR model, we do not include the state variables determining the processes of the technology and assets. The wealth is accumulated until the terminal period. At the end, the representative agent exhausts all wealth in consumption without bequest. This setting can be understood as a scenario where one accumulates wealth before retirement, and buys a pension product at retirement. Equivalently, this problem can be considered as a terminal wealth problem. The utility is a logarithm function, following Gagliardini et al. [23], who suggest this choice for its convenience to concentrate on the intuition of the impact of model uncertainty on the results whilst preserving tractability.

In the presence of model uncertainty, the representative agent evaluates the expected utility over a set of probability distributions, rather than a single nominal one. This set is the so-called model uncertainty set. The investment decision is made based on the probability leading to the worst utility expectation, reflecting a conservative attitude towards investment. There are many ways to formulate and solve such a problem; see the review by Guidolin and Rinaldi [29]. In this paper, we incorporate model uncertainty based on the foundation given by Hansen et al. [41]. It is shown in their paper that we are able to obtain time consistent solution for an intertemporal decision, under the formulation structure using a state and time independent "uncertainty parameter" .

Generally, the formulation in Hansen et al. [41] uses an equivalent set of Radon-Nikodym (RN) derivatives to represent the model uncertainty set. Since the RN derivative is defined as the ratio of an alternative probability and the nominal one, one can evaluate an alternative expectation by simply multiplying the corresponding RN derivative to the nominal model. In other words, the optimization can be conducted under the nominal measure, using the the RN derivative set as the constraint.

While this formulation uses the RN derivative explicitly, most related literature studying financial models under model uncertainty uses the RN derivatives implicitly. These formulations typically produce recursive results depending on a dynamic ”uncertainty process”. However, with the formulation explicitly using RN derivative, we obtain results depending on a Lagrangian-like ”uncertainty parameter”. To do so, if we simply follow Hansen et al. [41], it could result in a commitment issue, which means that the decision maker has to be committed to the initial plan, and cannot re-plan even if more information flow in future could change the optimal solution. Hansen and Sargent [34] discuss this type of problem, and propose to use so-called "discount" RN derivative, which is a modified RN derivative discounted by time, such that the uncertainty parameter is time invariant. However, we regard using "discount" RN derivative could reduce tractability.

Glasserman and Xu [27] apply Hansen et al. [41]'s formulation to study robust risk measure. Although their problem is not intertemporal, we find it inspiring. We propose in this paper to use the ordinary RN derivative, following their approach to obtain a state dependent and time consistent uncertainty parameter. However, regarding the issue of commitment and in order to maintain the time consistent property, we distinguish planning time and implementation time for convenience, and time consistency means implementation time consistency. In the new approach, we allow the amount of model uncertainty taken into account to be planning-time-varying, such that a (discount) uncertainty parameter is implementation-time-invariant. This is what makes our approach different from Hansen and Sargent [34]'s, while contributing to solve problems with the type of formulation by Hansen et al. [41].

Assuming a complete market, we solve the dynamic investment problem by the martingale approach. The results are compared with those in the absence of model uncertainty. We establish and analyze the links between the formulations implicitly embedding the RN derivative and those applied in this paper. We focus on two aspects. One is the link between the static uncertainty parameter and the dynamic uncertainty process, while the other is the risk premium and the market price of risk under different measures, based on the link in the first aspect.

Moreover, we investigate the risk-free rate, the risky rates, and the risk premium in equilibrium. The equilibrium is defined as in the original CIR, i.e., a representative agent finds an optimal strategy, given the market clearing condition that all wealth is invested in the production of the final consumption, and no wealth is invested in financial assets. We compare the terminal wealth in equilibrium, particularly between formulations under different measures.

We also calibrate the equilibrium solutions, using data of the OECD economic growth rates as the production technology returns. We consider two possibilities for the risky assets. One is a proxy of the market portfolio, and the other are five industry portfolios, both of which are from Kenneth French's data collection.[1] The data range from 2006 to 2017. Our calibrations consider the dynamic decision made over 1 year, 5 years, 10 years, and 15 years. At last, we investigate a reasonable amount of model uncertainty by a statistic in Pardo [61] that is related to the RN derivative. We classify model uncertainty into parameter uncertainty and misspecification uncertainty. The former can be eliminated given sufficient data, and appears between statistically equivalent (indistinguishable) models. The latter exists between models that are statistically indistinguishable, and cannot be eliminated only by increasing the sample size. The popular detection error probability approach in Anderson et al. [2] infers parameter uncertainty only, while using the statistic in Pardo [61] does not have such a restriction. In the calibrations, we find that the misspecification uncertainty is also needed to explain the equity risk premium puzzle, unlike in Maenhout [57] who finds parameter uncertainty is sufficient. The equilibrium risk-free rate, equity risk premium, and terminal wealth are analyzed in the calibrations.

The strand of literature on uncertainty, or ambiguity, dates back at least to "Knightian uncertainty" [49], which is distinct from the concept of risk, and applied to events with unknown probability.[2] The first famous related experiment is pro-

[1] Source: Kenneth French's website.

[2] Ambiguity, a fairly similar concept to uncertainty, is also discussed by Knight [49]. In this

vided by Ellsberg [20], who shows that the conventional subjective expected utility (SEU) is unable to account for the probability uncertainty; if SEU is used, the underlying probabilities implied in the experiment are inconsistent. This is called the "Ellsberg paradox". The experiment is also an evidence of ambiguity aversion, which refers to the preference of a choice with known probability to one with unknown probability.

The most popular proposal to settle the inconsistency in the Ellsberg paradox is given by Gilboa and Schmeidler [26]. They propose that in the presence of ambiguity, the decision maker (DM) considers a set of probability distributions, and makes decisions based on the probabilities leading to the worst expected outcomes. This is the max-min expected preference theory. The set of probabilities is typically referred to as the uncertainty set, containing plausible probabilities. The max-min rule is built on a conservative point of view that Nature always behaves against the DM. If the worst case is optimized, the DM would be well prepared for the other cases. This type of preferences is commonly known as multiple prior preferences (MPP).

Derived from Gilboa and Schmeidler [26]'s theory, there are two main streams of approaches for recursive intertemporal applications. One approach is led by Epstein and Schneider [21], who follow closely the MPP in the formulation of the problem to be investigated, and develop recursive MPP. The other approach is led by Hansen and Sargent [33], which combines the MPP with robustness optimization theory in engineering. Their main difference lies in the ways to form the constraint on the set of plausible probabilities, or the so-called uncertainty set. The former approach in fact contains multiple constraints on the uncertainty set, which are state and time dependent, and satisfy the rectangularity condition given by Chen and Epstein [8]. The latter approach takes into account only a single constraint on the uncertainty set, which is fixed and state independent, according to Anderson et al. [2] and Hansen et al. [41].

There have been a number of papers applying these two approaches to incorporate model uncertainty for a variety of topics in portfolio rules and asset pric-

paper, we use them interchangeably.

ing. To name a few, Maenhout [57] makes a modification on Anderson et al. [2] to allow for multiple constraints on the uncertainty set, in order to guarantee the homotheticity property in the portfolio problem. Trojani and Vanini [68] investigate the general equilibrium with Epstein-Zin preferences in the framework of Anderson et al. [2], but additionally incorporate the rectangularity condition for dynamic consistency, resulting in state-dependent constraints as well. A similar approach is applied by Gagliardini et al. [24] who study the CIR term structure of interest rates. Garlappi et al. [25] incorporate model uncertainty to Markowitz's mean-variance portfolio allocation, and propose a multi-prior approach. Ju and Miao [47], Liu [53], and others elaborate their problems with more financial setups, such as time-varying investment opportunities, regime-switching processes, and so on. However, all these papers apply multiple constraints to the uncertainty set based on either Chen and Epstein [8] or a modification of Anderson et al. [2], mainly because of the controversy on time consistency.

Later on, this controversy is solved as Maccheroni et al. [55, 56] prove that the original Hansen and Sargent [33] and Anderson et al. [2] are time consistent for intertemporal problems, and Hansen et al. [41] provide the axiomatic foundation for a single fixed and state independent constraint. Yet, we find limited literature applying such kind of constraint. Glasserman and Xu [27] apply such a constraint, and develop a more intuitive and practical approach for portfolio risk measurement, credit risk, hedging, etc. However, it does not study intertemporal decisions. Hansen and Sargent [34] and Hansen and Sargent [35] extend to discuss the issue under commitment and without commitment. Both of them propose using a discount RN derivative, which effectively shrinks the uncertainty set over time, to guarantee a time consistent solution. Our paper, as mentioned earlier, proposes to use a planning-time-varying amount of model uncertainty instead, and combine Glasserman and Xu [27] 's approach, resulting in a much simpler method.

In terms of topics and settings, the closely related literature considering the logarithm utility or the terminal wealth under model uncertainty includes Maenhout [57], and Gagliardini et al. [23]. The former derives the solutions for a logarithm utility from a general setting of a CRRA utility, where the consumption occurs every period; our problem considers a modified CIR problem with only terminal consumption. For a terminal wealth problem, the later imposes multiple constraints for time consistency, while we use a single constraint.

This paper contributes to the literature studying portfolio rules and asset pricing under model uncertainty. Firstly, in terms of methodology, we propose an alternative approach to Hansen and Sargent [34] to obtain a fixed and state independent (discount) uncertainty parameter, while maintaining time consistency under commitment. Secondly, we find that the demand for the risk-free asset increases, while the demand for production improvement and risky assets decrease. This finding supports Maenhout [58], but not Liu [53] who finds increasing demand for risky assets when the elasticity of intertemporal substitution is taken into account. Also, we find that the conservative attitude of the representative agent in equilibrium is affected not just by the amount of model uncertainty, but also by the investment horizon and the volatilities of production improvement. Last but not least, our calibrations show that the risk-free rate and equity risk premium puzzle can be relaxed by incorporating model uncertainty. Unlike Maenhout [57], we suggest additional model misspecification uncertainty to unveil unexplained premium.

In the following sections, Section 4.2 first makes a brief mathematical review on the mainstream representations and formulations to incorporate model uncertainty. It then proceeds to describe the general financial market, incorporate model uncertainty, and solve the investment-consumption problem under model uncertainty. Section 4.3 discusses the implication of the solutions. Section 4.4 analyzes specifically the equilibrium case. Section 4.5 provides calibration results to understand the impact of the crucial uncertainty parameter in equilibrium. Section 4.6 summarizes and concludes. Appendix 4.A provides the assumptions for our main formulation. We provide the details in solving the optimization in Appendix 4.B, and the equilibrium in Appendix 4.C.

4.2 The Investor's Optimization Problem

Financial models are set up to explain the relations between financial quantities such as investment, consumption, risk-free rates, pricing factors, discounted rates, etc. The classical models typically consider a nominal probability distribution, denoted by $\mathbb{P}$. However, there are quite many reasons why the true outcomes in the economy do not perform as estimated and predicted by the nominal model $\mathbb{P}$. These reasons may be caused by the estimation approach, model structure settings, or even unexpected changes in reality. Under such circumstances, the outcome is actually derived from an alternative probability distribution $\widetilde{\mathbb{P}}$. The discrepancy between $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ raises model uncertainty. This section incorporates model uncertainty into a modified CIR problem, and derives closed-form solutions.

We will first briefly review some possible formulations for model uncertainty incorporation. Then we describe the financial market settings, the nominal optimization problem. The last part shows the incorporation of model uncertainty, and the closed-form solutions.

4.2.1 A Brief Review of Some Alternative Formulations

This part provides a not-so-formal review on the main formulations pertaining to this paper. We write $x \sim \widetilde{\mathbb{P}}$, where x is a random variable of interest, distributed according to a probability distribution $\widetilde{\mathbb{P}}$ in an uncertainty set $\mathcal{P}$. According to Hansen and Sargent [33], we could formulate the investor's preferences in two ways to incorporate model uncertainty, as follows,

$$(P1) \qquad \inf_{\widetilde{\mathbb{P}} \in \widetilde{\mathcal{P}}(x)} \int U(x) d\widetilde{\mathbb{P}}, \quad \text{subject to} \quad \mathcal{R}(\widetilde{\mathbb{P}}) \leq \kappa,$$

$$(P2) \qquad \inf_{\widetilde{\mathbb{P}} \in \widetilde{\mathcal{P}}(x)} \int U(x) d\widetilde{\mathbb{P}} + \theta \mathcal{R}(\widetilde{\mathbb{P}}),$$

where $U(x)$ is the utility function of x. The parameter κ represents the amount of model uncertainty as specified by the decision maker (hereinafter, DM). $\mathcal{R}(\widetilde{\mathbb{P}})$ is the relative entropy, or the Kullback-Leibler divergence (hereinafter, KL divergence) [50]. It measures the discrepancy between the distribution $\widetilde{\mathbb{P}}$ and the nominal distribution $\mathbb{P}$, quantifying model uncertainty.[3] Specifically, the KL divergence is defined by $\mathcal{R}(\widetilde{\mathbb{P}}) = \mathbb{E}^{\widetilde{\mathbb{P}}}(\log m)$, where $m = d\widetilde{\mathbb{P}}/d\mathbb{P}$ is a well-defined likelihood ratio called the Radon-Nikodym derivative (hereinafter, RN derivative). In a continuous intertemporal context, the RN derivative is a martingale process, with the property

$$m_t = \mathbb{E}\left(m_T \middle| \mathcal{F}_t\right), \qquad \forall t < T, \tag{4.1}$$

where $\mathcal{F}_t = \sigma\{Z_s, 0 \le s \le t\}$ for $t \in [0, T]$ is the filtration generated by Z_t, which we assume to be a Brownian motion under the measure $\mathbb{P}$. The filtration $\mathcal{F}_t$ collects historical information. Moreover, by Girsanov Theorem,

$$m_t = m_0 \exp\left(-\frac{1}{2}\int_0^t h_s' h_s ds + \int_0^t h_s' dZ_s\right), \tag{4.2}$$

where $m_0 = 1$, and h_t is a process adapted to $\mathcal{F}_t$.

The formulation $(P1)$ is usually called the constraint problem, while $(P2)$ the penalty problem because the objective includes an extra term related to the constraint $\widetilde{\mathbb{P}} \in \widetilde{\mathcal{P}}(x)$. The parameter θ is an implied Lagrangian multiplier on the constraint in $(P1)$. Because it reflects the impact of model uncertainty, we call it the "uncertainty parameter". The larger θ is, the smaller the impact of model uncertainty is. Mostly for the convenience of interpretation, we use $\frac{1}{\theta}$ instead of θ. According to Hansen and Sargent [33], the orderings of two preference formulations are not necessarily the same, but optimizing over them provides identical decisions.

Alternatively, with the help of m_t, $(P1)$ and $(P2)$ can be rewritten in terms of $\mathbb{P}$, i.e,

$$(P1') \qquad \inf_{\mathbb{P}\in\mathcal{P}} \int m(x)\cdot U(x)d\mathbb{P}(x), \quad \text{subject to} \quad \mathcal{R}(\mathbb{P}) \le \kappa,$$

$$(P2') \qquad \inf_{\mathbb{P}\in\mathcal{P}} \int m(x)\cdot U(x)d\mathbb{P}(x) + \theta\mathcal{R}(\mathbb{P}),$$

[3] There are other types of f-divergences to quantify model uncertainty. The KL divergence is the one mostly used in practice, and applicable for our problem.

where $\mathcal{R}(\mathbb{P}) = \mathbb{E}^{\mathbb{P}}(m \log m)$.

A crucial part in the problems is constructing the uncertainty set, or formulating the constraint(s), for the intertemporal investigation. There are two kinds of constructions, based on [8] and Hansen et al. [41], respectively. Both of them can be postulated through h_t,

$$\begin{aligned} &(A) \qquad \{h : h_t' h_t \leq \kappa\}, \\ &(B) \qquad \left\{h : \mathbb{E}^{\mathbb{P}}(m_T \log m_T) \leq \kappa, \text{ and } m_T \text{ defined by } (4.2)\right\}. \end{aligned}$$

The main difference between these two constraints is that (A) implies multiple state-dependent and time-dependent constraints; see Trojani and Vanini [68]. They satisfy the rectangularity condition described in Chen and Epstein [8], according to Trojani and Vanini [68]. It is the type used in [24, 57, 58, 68]. Because h_t is the essential process reflecting the size of the uncertainty set, and determining m_t, we call it the "uncertainty process".

The constraint (B), instead, implies only a single constraint related to the terminal RN derivative m_T, yielding a fixed and state independent "uncertainty parameter" θ. In fact, (B) applies the martingale property of m_t, and needs only the terminal RN derivative m_T. Under this constraint formulation, the optimization problem need not be solved recursively, and thus raise the controversy of time consistency. Maccheroni et al. [55, 56] show that it suffices to use this single constraint to maintain recursive consistency. However, in order to maintain time consistency, Hansen et al. [41] use a so-called discount RN derivative $\widetilde{m}_t$, instead of the ordinary m_t, to establish the solution approach axiomatically. Hansen and Sargent [34] and Hansen and Sargent [35] discuss the control under commitment and without commitment, respectively.

Our paper applies $(P1')$. But different from Hansen et al. [41], we remain using the ordinary RN derivative. However, in order to guarantee time consistent solution under commitment, we clarify planning time τ and implementation time t. We allow the amount of model uncertainty κ to be planning-time-varying, depending on the planning time τ. When $\tau = 0$, $\kappa_\tau = \kappa$. The constraint (B) is now,

$$(B') \qquad \left\{h : \mathbb{E}^{\mathbb{P}}\left(m_T \log m_T\right) \leq \kappa_\tau, \text{ and } m_T \text{ defined in the form } (4.2)^4\right\},$$

In other words, we will formulate a modified Cox et al. [11] investment-consumption problem the way as $(P1')$ and (B'), and study the impact of model uncertainty. Unless particularly specified, the RN derivative in the sequel refers to the ordinary one.

4.2.2 The Nominal Model of the Financial Market

The financial market follows the settings described in Cox et al. [11], except for taking into account the state variables determining other parameters in the market components. We also assume a complete market. The governing measure $\mathbb{P}$ is reflected in the standard Brownian motions driving the diffusion parts in all processes, collected in the matrix $Z_t \in \mathbb{R}^{(1+k)\times 1}$, where k is the number of risky assets. Specifically, we have the following components in the market.

1. There is one risk-free asset (fixed-income security) B_t with return r_t,

$$dB_t = r_t B_t dt.$$

2. The output rate of a linear production technology Q_t is postulated as

$$dQ_t = \alpha_{1t} Q_t dt + \sigma_{1t} Q_t dZ_t.$$

   where $\alpha_{1t} \in \mathbb{R}^{1\times 1}$ is the non-random drift term of the process, and $\sigma_{1t} \in \mathbb{R}^{1\times(1+k)}$ is a non-random row vector. The parameter α_{1t} is in fact the expected return at time t. A full rank σ_{1t} indicates that Q_t is affected by other diffusion components. The Q_t process produces only *one* physical good,

[4] Equation (4.2) is defined when $\tau = 0$. When introducing τ, m_T is defined by $m_t = m_\tau \exp\left(-\frac{1}{2}\int_\tau^t h_s' h_s ds + \int_\tau^t h_s' dZ_s\right)$ for $\tau < t$, where $m_\tau = 1$.

which can be either reinvested or consumed. This postulated process represents the production side of the economy. The production is used as a numéraire; all values are expressed in terms of the units of this good.

3. The market provides k risky assets, complied in the price vector $S_t = \left[S_{1t}, \cdots S_{kt}\right]'$, evolving according to

$$dS_t = I_{S_t} \alpha_{2t} dt + I_{S_t} \sigma_{2t} dZ_t,$$

where $I_{S_t} = diag\left(S_t\right)$. α_{2t} is a non-random vector collecting the drift terms of the risky assets, and $\sigma_{2t} \in \mathbb{R}^{k\times(k+1)}$ is a non-random matrix describing the volatilities of the risky assets and the single product.

For notational convenience, denote

$$\alpha_t = \begin{bmatrix} \alpha_{1t} \\ \alpha_{2t} \end{bmatrix}, \qquad \Sigma_t = \begin{bmatrix} \sigma_{1t} \\ \sigma_{2t} \end{bmatrix}, \qquad \pi_t = \begin{bmatrix} \pi_{1t} \\ \pi_{2t} \end{bmatrix},$$

where π_{1t} and π_{2t} are the portions of wealth invested in production and risky assets, respectively. The vector α_t collects the drift terms, which are in fact the expected returns at time t. The matrix Σ_t is assumed full rank, to satisfy the completeness of the market. The variance-covariance of the single product and risky assets is $\Sigma_t \Sigma_t'$.[5]

The representative agent invests every period without consuming. The wealth is accumulated to be W_T at the terminal period and is exhausted in consumption. The utility function is a logarithm function, i.e., $U(C_T) = \log\left(C_T\right)$, where $C_T = W_T$. The choice of logarithm as the utility function follows Gagliardini et al. [23], who argue that the logarithm setting is convenient to focus on the intuition of model uncertainty while being highly tractable. The condition $C_T = W_T$ is interesting for a pensioner, who accumulates wealth before retirement, experiences economic

[5] The variance-covariance of the single product and risky assets is estimated by $cov\left(\begin{bmatrix} \sigma_{1t} dZ_t \\ \sigma_{2t} dZ_t \end{bmatrix} \begin{bmatrix} \sigma_{1t} dZ_t & \sigma_{2t} dZ_t \end{bmatrix}\right) = \begin{bmatrix} \sigma_{1t}\sigma_{1t}' & \sigma_{1t}\sigma_{2t}' \\ \sigma_{2t}\sigma_{1t}' & \sigma_{2t}\sigma_{2t}' \end{bmatrix} dt$. Thus σ_{1t} and σ_{2t} can be constructed such that this condition is satisfied, by Cholesky decomposition of $\Sigma_t \Sigma_t'$ for instance.

growth by technology improvement, and spends all wealth on a pension product in the last period.

Suppose b is a deterministic discount rate. In the absence of model uncertainty, the representative agent maximizes the expectation of the subjectively discounted utility under the measure $\mathbb{P}$

$$\max_{\pi_t} \mathbb{E}^{\mathbb{P}} \left[\exp\left(-bT\right) \log\left(C_T\right) \right], \qquad C_T > 0, \tag{4.3}$$

subject to the terminal condition

$$C_T = W_T, \tag{4.4}$$

and the wealth dynamics under $\mathbb{P}$, postulated as

$$\frac{dW_t}{W_t} = \left[\pi_t' \left(\alpha_t - r_t \mathbf{1}_{k+1} \right) + r_t \right] dt + \pi_t' \Sigma_t dZ_t, \tag{4.5}$$

where $\mathbf{1}_i$ is a $\mathbb{R}^{i\times 1}$ vector of ones.

In this nominal setting, the analyses of investment decision, interest rates, risk premium, etc. are based on this single prior measure $\mathbb{P}$.

4.2.3 The Optimization Problem under Model Uncertainty

It is crucial to construct a model uncertainty set that implies a specific amount of model uncertainty. In this paper, the amount of model uncertainty is quantified through the KL divergence, which is defined by the RN derivative. Meanwhile, an alternative model can be represented through a change of measure, i.e., by the RN derivative as well. Because of these links with the RN derivative, the uncertainty set containing alternative models is equivalent to an uncertainty set containing the RN derivatives, whose corresponding KL divergences satisfy the constraint for the specified amount of model uncertainty. In other words, we can simply apply the uncertainty set with respect to the RN derivative, which is more practical to handle.

Having in mind all possible models in the uncertainty set, a conservative agent cares most about the one leading to the worst expected outcome. He aims to optimize the worst case, because once the worst case is treated optimally, he could be relieved from worrying about the other cases. In other words, the optimization problem is to maximize the worst of all expectations given by the alternative models. Moreover, we apply constraint (B') here, and distinguish planning time τ and implementation time t. Thus, the optimization problem is given by [6]

$$\sup_{\pi_t} \inf_{m_T \in \mathcal{P}_{\kappa_\tau}} \mathbb{E}^{\mathbb{P}} \left\{ m_T \exp[-b(T-\tau)] \log(W_T) \right\}, \tag{4.6}$$

where m_T is the RN derivative that changes the measure when evaluating the utility expectation at the terminal period T, and τ is the time making the optimal plan. The set $\mathcal{P}_{\kappa_\tau} = \left\{ m_T | \, \mathbb{E}^{\mathbb{P}} \left(m_T \log m_T \right) \leq \kappa_\tau \right\}$ collects all possible m_T that satisfy the constraint on the specific amount of model uncertainty at time τ, κ_τ. Since an m_T corresponds to an alternative model changed from the nominal one, the collections of m_T, $\mathcal{P}_{\kappa_\tau}$, is equivalent to a model uncertainty set.

The inner infimum in (4.6) reflects the agent's concern on the worst outcome given by different models. His optimal investment decision is made so as to maximize the worst case. The belief is that if the investment could gain in the worst case, so would it do in other cases. Nevertheless, the worries on the worst case restrict his investment, and thus his behavior is regarded as conservative.

4.2.4 The Solutions to the Optimization Problem

In the sequel, the wealth W_T is subject to the dynamic process (4.5). Following Glasserman and Xu [27], the inner infimum problem in (4.6) can be reformulated into a Lagrangian dual function, associated with a Lagrangian multiplier $\frac{1}{\theta_\tau}$ for the constraint at planning time τ. Consequently, the optimization problem becomes

[6] Using the constraint type (B'), it was $\sup_{\pi_{t,\tau}} \inf_{m_t \in \mathcal{P}_{\kappa_\tau}} \mathbb{E}^{\mathbb{P}} \left[m_t U(\pi_{t,\tau}) \right]$. Applying the martingale property of m_t process, it becomes $\sup_{\pi_{t,\tau}} \inf_{m_T \in \mathcal{P}_{\kappa_\tau}} \mathbb{E}^{\mathbb{P}} \left[m_T U(\pi_{t,\tau}) \right]$, where m_T aggregates the historical information. It is sufficient to use a single constraint with respect to the terminal value of the RN derivative process m_T.

[7]

$$\sup_{\pi_{t,\tau}} \sup_{\theta_\tau<0} \inf_{m_T} \left\{ \mathbb{E}^{\mathbb{P}} \left\{ m_T \exp[-b(T-\tau)] \log(W_T) \right\} - \frac{1}{\theta_\tau} \mathbb{E}^{\mathbb{P}} \left(m_T \log m_T - \kappa_\tau \right) \right\}. \tag{4.7}$$

Under appropriate assumptions satisfied in our case (see Appendix 4.A), and following Proposition 2.1 in Glasserman and Xu [27], we are allowed to interchange the optimizers $\sup_{\pi_{t,\tau}}$ and $\sup_{\theta_\tau<0}$,

$$\sup_{\theta_\tau<0} \sup_{\pi_{t,\tau}} \inf_{m_T} \mathbb{E}^{\mathbb{P}} \left\{ m_T \exp[-b(T-\tau)] \log\left(W_T\right) - \frac{1}{\theta_\tau}(m_T \log m_T - \kappa_\tau) \right\}. \tag{4.8}$$

This problem is solved from the most inside layer. The optimizer $\inf_{m_T}$ solves for the optimal m_T, producing a θ_τ-associated optimal solution

$$m_T^*(\theta_\tau) = \frac{\exp\left\{\theta_\tau \exp[-b(T-\tau)] \log\left(W_T^*\right)\right\}}{\mathbb{E}^{\mathbb{P}}\left\{\exp\left\{\theta_\tau \exp[-b(T-\tau)] \log\left(W_T^*\right)\right\}\right\}}, \quad \theta_\tau < 0, \tag{4.9}$$

where W_T^* is the terminal wealth, depending on the optimal investment decision $\pi^*_{t,\tau}$. By substituting $m_T^*(\theta_\tau)$ into (4.8), the optimal value for second layer optimizer $\sup_{\pi_{t,\tau}}$ can be defined by

$$J(\theta_\tau) := \frac{1}{\theta_\tau} \log \mathbb{E}^{\mathbb{P}} \left\{ \exp\left\{\theta_\tau \exp[-b(T-\tau)] \log\left(W_T^*\right)\right\}\right\},$$

---

[7] The problem (4.7) is derived from the dual problem of (4.6)

$$\sup_{\pi_{t,\tau}} \sup_{\tilde\theta_\tau>0} \inf_{m_T} \left\{ \mathbb{E}^{\mathbb{P}} \left\{ m_T \exp[-b(T-\tau)] \log(W_T) \right\} + \frac{1}{\tilde\theta_\tau} \mathbb{E}^{\mathbb{P}} \left(m_T \log m_T - \kappa_\tau \right) \right\}.$$

Let $\theta_\tau = -\tilde\theta_\tau$. It then becomes (4.7). The problem (4.7) can also be changed to $\inf_{\pi_t} \inf_{\theta_\tau<0} \sup_{m_T} \left\{ -\mathbb{E}^{\mathbb{P}} \left\{ m_T \exp[-b(T-\tau)] \log(W_T) \right\} + \frac{1}{\theta_\tau} \mathbb{E}^{\mathbb{P}} \left(m_T \log m_T - \kappa_\tau \right) \right\}$, and further

$$\inf_{\pi_{t,\tau}} \inf_{\theta_\tau>0} \sup_{m_T} \left\{ -\mathbb{E}^{\mathbb{P}} \left\{ m_T \exp[-b(T-\tau)] \log(W_T) \right\} - \frac{1}{\theta_\tau} \mathbb{E}^{\mathbb{P}} \left(m_T \log m_T - \kappa_\tau \right) \right\},$$

which in line with the Proposition 2.1 in Glasserman and Xu [27] that studies a minimization problem of a convex objective. Then we are allowed to apply the proposition in Glasserman and Xu [27].

where W_T^* results from optimal investment decision $\pi_{t,\tau}^*$, and $\pi_{t,\tau}^*$ can be expressed as

$$\pi_{t,\tau}^* = \arg\sup_{\pi_{t,\tau}t} \frac{1}{\theta_\tau} \log \mathbb{E}^{\mathbb{P}} \left\{ \exp\left\{ \theta_\tau \exp[-b(T-\tau)] \log\left(W_T\right) \right\} \right\}.$$

For a given value of θ_τ concerning in the first layer optimization, the optimal value for the inner part $\sup_{\pi_{t,\tau}} \inf_{m_T}$ is in fact give by

$$J(\theta_\tau) = \sup_{\pi_{t,\tau}} \frac{1}{\theta_\tau} \log \mathbb{E}^{\mathbb{P}} \left\{ \exp\left\{ \theta_\tau \exp[-b(T-\tau)] \log\left(W_T\right) \right\} \right\},$$

resulting from the optima $\pi_{t,\tau}^*$ and m_T^*. As a consequence, the optimization (4.8) can be reduced to a problem with respect to θ_τ,

$$\sup_{\theta_\tau < 0} \left[J(\theta_\tau) + \frac{\kappa_\tau}{\theta_\tau} \right] = \sup_{\theta_\tau < 0} \sup_{\pi_{t,\tau}} \left\{ \frac{1}{\theta_\tau} \log \mathbb{E}^{\mathbb{P}} \left\{ \exp\left\{ \theta_\tau \exp[-b(T-\tau)] \log\left(W_T\right) \right\} \right\} + \frac{\kappa_\tau}{\theta_\tau} \right\}. \tag{4.10}$$

We start from solving the inner optimization problem with respect to $\pi_{t,\tau}$ in (4.10), and then proceed further to find the optimum θ_τ^*. Unlike Cox et al. [11] and other literature to follow Merton [59], we will solve the inner optimization problem in (4.10) by the martingale approach, following Karatzas and Shreve [48] and making use of the assumption of a complete market. The advantage of employing the martingale approach is that it avoids the complicated optimization steps in the recursive dynamic method in Merton [59]. The disadvantage is that it is restricted to the complete market setting. In an incomplete market, the Hamilton-Jacobian-Bellman equation has to be used to ensure the recursion consistency for intertemporal decisions.

For notational simplicity, write $a_{\theta_\tau} = \theta_\tau X_{T,\tau}$ and $X_{T,\tau} = \exp[-b(T-\tau)]$ for a given θ_τ. The inner supremum problem under the measure $\mathbb{P}$ is equivalent to [8]

$$\sup_{\pi_{t,\tau}} \left[-\mathbb{E}\left(W_T^{a_{\theta_\tau}}\right)\right],$$

subject to (4.5). As $X_{T,\tau}$ is a discount factor, one may think of a_{θ_τ} as the uncertainty parameter θ_τ discounted by $X_{T,\tau}$.

We show the solution details in Appendix 4.B. Moreover, the outer optimization for the optimal θ^*_τ is easily implemented when $\pi_{t,\tau}$ is found; see also Appendix 4.B. The solutions are summarized in the following proposition.

Proposition 2. *Planned at time τ, the optimal investment decision for period $t \geq \tau$, is given by*

$$\pi^*_{t,\tau} = \frac{1}{1 - a_{\theta^*_\tau}} (\Sigma_t \Sigma_t')^{-1} (\alpha_t - r_t \mathbf{1}_{k+1}), \qquad t \geq \tau, \tag{4.11}$$

and the terminal wealth yields

$$\begin{aligned} W^*_T = & W_\tau \exp\left\{ \int_\tau^T \left[\frac{1 - 2a_{\theta^*_\tau}}{2(1 - a_{\theta^*_\tau})^2} (\alpha_t - r_t \mathbf{1}_{k+1})' (\Sigma_t \Sigma_t')^{-1} (\alpha_t - r_t \mathbf{1}_{k+1}) + r_t \right] dt \right. \\ & \left. + \int_\tau^T \frac{1}{1 - a_{\theta^*_\tau}} (\alpha_t - r_t \mathbf{1}_{k+1})' \Sigma_t'^{-1} dZ_t \right\}, \end{aligned} \tag{4.12}$$

where $a_{\theta^*_\tau} = \theta^*_\tau \exp\left[-b\left(T - \tau\right)\right]$ with the optimal θ^*_τ given by

$$\theta^*_\tau = \frac{-\kappa_\tau - \sqrt{\kappa_\tau \int_\tau^T (\alpha_t - r_t \mathbf{1}_{k+1})' (\Sigma_t \Sigma_t')^{-1} (\alpha_t - r_t \mathbf{1}_{k+1}) dt}}{\left[\int_\tau^T (\alpha_t - r_t \mathbf{1}_{k+1})' (\Sigma_t \Sigma_t')^{-1} (\alpha_t - r_t \mathbf{1}_{k+1}) dt - \kappa_\tau\right] \exp\left[-b\left(T - \tau\right)\right]} < 0, \tag{4.13}$$

for an amount of model uncertainty $\kappa_\tau > 0$ at the planning time τ.

[8] The original problem can be written as $\sup_{\pi_{t,\tau}} \frac{1}{\theta_\tau} \log \mathbb{E}\left(W_T^{a_{\theta_\tau}}\right)$. The part $\frac{1}{\theta_\tau}$ is not relevant. The log function is monotonically increasing. But $a_{\theta_\tau} < 0$ makes the objective function a monotonically decreasing one. The original problem is thus equivalent to optimizing a monotonically increasing function of W_T by adding a negative sign.

Under this formulation, the investment decision $\pi^*_{t,\tau}$ depends also on τ. This means that the optimal solution depends on the initial state at planning, accompanied with a commitment issue. If we need implementation-time consistency, the planning time τ should not affect. In other words, $a_{\theta^*_\tau}$ should be consistent such that $\pi^*_{t,\tau} = \pi^*_t$. Using (4.13) for $a_{\theta^*_\tau}$, both the amount of model uncertainty and the integration part depend on τ. Therefore, one way to achieve time consistency is to allow the amount of model uncertainty κ_τ to be planning-time-varying, such that for each planning time τ, we could find an amount κ_τ resulting in a consistent and fixed $a_{\theta^*_\tau}$. Since $a_{\theta^*_\tau}$ is consistent regardless of τ, we can obtain the fixed a_{θ^*} by simply considering the initial state of amount of model uncertainty κ_0. Consequently, the implementation-time consistent investment strategy π^*_t is established.

In contrast to our approach, Hansen and Sargent [34] keep κ_τ consistent, but using a discounted RN derivative instead of an ordinary one. This could reduce the tractability in practice. Although our approach seem to have complicated structure κ_τ with extra planning time τ, the solution can be obtained by the initial state of model uncertainty. The feature makes it attractive in the solving process. In the sequel, κ is used instead of κ_τ for the initial model uncertainty state.

With time consistency, Proposition 2 states the optimal investment decision π^*_t, the terminal wealth W^*_T under the measure $\mathbb{P}$. The discounted uncertainty parameter a_{θ^*} is in fact not related to the initial wealth W_τ, and the discount rate b (because of the factor $1/\exp[-b(T-\tau)]$ in θ^*_τ). It decreases in κ, and is negative when the integration in the denominator of θ^*_τ is larger than κ.[9] It is time-invariant, determined by κ and an integration function over the length of the time horizon T with the expected investment returns a_t, the risk-free rate r_t, and the innovations between investments Σ_t. If initial model uncertainty approaches to 0, i.e., $\kappa \to 0$, the optimization problem gets close to the nominal one. In the limit, the solutions are simply the nominal ones, as if $a_{\theta^*} = 0$.

The investment decision π^*_t depends on model uncertainty by a_{θ^*}. A more conservative agent considers more model uncertainty, i.e., a huger κ^*. As κ^* gets larger, a_{θ^*} gets smaller since it is a negative parameter, and hence the coefficient $\frac{1}{1-a_{\theta^*}}$ is

[9] This is most likely to be the case, and will be discussed by calibration in Section 4.5.

smaller. This means that the demand to invest in production improvement and risky assets is less, while the demand of lending or borrowing the risk-free asset increases. This supports the finding in Maenhout [58]. The investor shrinks the risky investment willingness in the presence of more model uncertainty.

Moreover, the investment decision π_t^* can be seen as an effective increase in the risk aversion by $-a_\theta^*$. This is in line with the findings in Liu [53], Maenhout [57, 58], and echoes the empirical findings that a high risk aversion coefficient can be reconciled with a preference to robustness; see e.g. Vissing-Jørgensen and Attanasio [70], and Bansal and Yaron [3].

The optimal terminal wealth is random, resulting from the Brownian motion Z_t. More model uncertainty means that the diffusion part is less influential because of the decreasing relation with $\frac{1}{1-a_{\theta^*}}$. Concerning the drift term, we have to consider that $1 > \frac{1}{1-a_{\theta^*}} = \frac{2-2a_{\theta^*}}{2(1-a_{\theta^*})^2} > \frac{1-2a_{\theta^*}}{2(1-a_{\theta^*})^2}$. Thus the optimal terminal wealth with model uncertainty, W_T^*, will be smaller than the one without model uncertainty, resulting from the conservative investment choice.

4.3 Implications

The way we incorporate model uncertainty is imposed on the nominal objective function (4.3) under the nominal measure $\mathbb{P}$, different from the framework designed by Hansen and Sargent [36]. The latter imposes model uncertainty on the constraint (4.5) by Girsanov's theorem. This changes the measure evaluating the nominal objective function to be an alternative one, $\widetilde{\mathbb{P}}$. In this section, we will compare the differences and the connections between these two formulations. Based on these connections, we develop a few implications.

4.3.1 The Alternative Formulation under $\widetilde{\mathbb{P}}$

In the formulation in Section 4.2, the RN derivative m_T is the most crucial element to transform the problem between measures. Hence, we make a further investigation on it, and explore how this transformation in $(P1')$ links to the formulation $(P1)$.

In particular, under $(P1)$, the optimization problem is solved recursively. Therefore, time-consistency is guaranteed, and commitment is not an issue. There is no need to distinguish the planning time τ. m_T is the end-point of the RN derivative process $m_t \in \mathbb{R}^+$, following

$$dm_t = m_t h_t' dZ_t,$$

or (4.2). When $t = T$, the end-period m_T is

$$m_T = m_0 \exp\left(\int_0^T h_t' dZ_t - \frac{1}{2}\int_0^T h_t' h_t dt\right). \tag{4.14}$$

If reconciled with the uncertainty amount κ, the RN derivative in (4.14) is equal to (4.9). Equating (4.9) and (4.14), we can establish the relation between the uncertainty parameter θ and the uncertainty process h_t.

Alternatively to making a change in the objective function, the classic way in Hansen and Sargent [36] to incorporate model uncertainty considers the model change in the wealth dynamic constraint, i.e,

$$\sup_{\pi_t} \inf_{h_t \in \mathcal{H}} \mathbb{E}^{\widetilde{\mathbb{P}}} \left[\exp(-bT) \log W_T\right] \tag{4.15}$$

subject to

$$\frac{dW_t}{W_t} = \left[\pi_t'(\alpha_t - r_t \mathbf{1}_{k+1} + \Sigma_t h_t) + r_t\right] dt + \pi_t' \Sigma_t d\widetilde{Z}_t,$$

where, by Gisanov's theorem, $d\widetilde{Z}_t = dZ_t - h_t dt$ is a standard Brownian motion under the alternative measure $\widetilde{\mathbb{P}}$ given by $\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = m_T$. One can think of this change as if the randomness driving force Z_t of the original measure $\mathbb{P}$ is shifted by a drift term value $-h_t dt$, and becomes $\widetilde{Z}_t$ under the alternative measure $\widetilde{\mathbb{P}}$. The uncertainty process h_t is considered within the constraint of a set $\mathcal{H}$ indicating the model uncertainty amount, analogous to the condition m_T in $\mathcal{P}_\kappa$ in (4.6). Through the uncertainty process h_t, m_t is implicitly embedded, and model uncertainty is incorporated into the wealth dynamics under the new measure $\widetilde{\mathbb{P}}$. The expected utility is hence evaluated with respect to $\widetilde{\mathbb{P}}$. In contrast, in the problem under measure $\mathbb{P}$, the uncertainty is incorporated only when evaluating the terminal expected utility. It is reflected in the investment decision only through the time-invariant discount uncertainty parameter a_θ, instead of the possibly complicated and time-varying process h_t. This also makes the (B') constraint more attractive in applications.

4.3.2 The Relation between a_θ and h_t

The relation between a_θ and h_t is pivotal to connect the measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$. Essentially, the key point lies in the property of the RN derivative process m_t.

Recall the normalized RN derivative in (4.9), where $m_T \propto \exp\left[a_\theta \log\left(W_T^*\right)\right]$ is normalized by the expectation $\mathbb{E}\left\{\exp\left[a_\theta \log\left(W_T^*\right)\right]\right\}$. It can be rewritten as $\exp\left(a_\theta \log W_T^* + \varrho\right)$, where $\varrho = -\mathbb{E}\left\{\exp\left[a_\theta \log\left(W_T^*\right)\right]\right\}$. Equating the expressions of the terminal RN derivative in (4.9) and (4.14), and taking out the exponential power of both sides, we have the following equality

$$a_\theta \log W_T^* - \log \mathbb{E}\left[\exp\left(a_\theta \log W_T^*\right)\right] = \int_0^T h_t' dZ_t - \int_0^T \frac{1}{2} h_t' h_t dt.$$

Substituting (4.12) into this equation results in

$$\begin{aligned}
&\int_0^T -\frac{a_\theta^2}{2\left(1-a_\theta\right)^2}\left(\alpha_t - r_t \mathbf{1}_{k+1}\right)' \left(\Sigma\Sigma'\right)^{-1}\left(\alpha_t - r_t \mathbf{1}_{k+1}\right) dt \\
&+ \int_0^T \frac{a_\theta}{1-a_\theta}\left(\alpha_t - r_t \mathbf{1}_{k+1}\right)' \Sigma'^{-1} dZ_t \\
= &\int_0^T h_t' dZ_t - \int_0^T \frac{1}{2} h_t' h_t dt.
\end{aligned} \tag{4.16}$$

Equating the diffusion parts on both sides of this equation implies that

$$h_t = \frac{a_\theta}{1-a_\theta} \Sigma_t^{-1}\left(\alpha_t - r_t \mathbf{1}_{k+1}\right) + A_t,$$

where $A_t \in \mathbb{R}^{(1+k)\times 1}$ is a random vector satisfying $A_t' Z_t = 0$. With this formula,

$$\begin{aligned} h_t' h_t = & \frac{a_\theta^2}{(1-a_\theta)^2} (\alpha_t - r_t \mathbf{1}_{k+1})' \left(\Sigma_t \Sigma_t'\right)^{-1} (\alpha_t - r_t \mathbf{1}_{k+1}) \\ & + \frac{2a_\theta}{1-a_\theta} A_t' \Sigma_t^{-1} (\alpha_t - r_t \mathbf{1}_{k+1}) + A_t' A_t. \end{aligned}$$

However, if and only if $A_t = 0$, h_t will be in line with the implication given by the drift terms on both sides of the equation (4.16), i.e.,

$$h_t' h_t = \frac{a_\theta^2}{(1-a_\theta)^2} (\alpha_t - r_t \mathbf{1}_{k+1})' \left(\Sigma_t \Sigma_t'\right)^{-1} (\alpha_t - r_t \mathbf{1}_{k+1}).$$

4.3.3 The Risk Premium and the Market Price of Risk

Under the measure $\widetilde{\mathbb{P}}$, we investigate the risk premium of the production technology and the risky assets, taking into account model uncertainty. Denote by R_t the vector stacking Q_t and S_t, i.e., $R_t = [Q_t, S_t']'$. Then, the process of R_t is given by

$$I_R^{-1} dR_t = (\alpha_t + \Sigma_t h_t)\, dt + \Sigma d\widetilde{Z}_t,$$

where $I_R^{-1} = diag\left(\frac{1}{Q_t}, \frac{1}{S_1}, \cdots, \frac{1}{S_k}\right)$.

The risk-free asset is unaffected by a change of measure, and thus the risk premium $d\widetilde{EP}$ with respect to $\widetilde{\mathbb{P}}$ is

$$\begin{aligned} d\widetilde{EP} = & \mathbb{E}\left(I_R^{-1} dR_t - B_t^{-1} \mathbf{1}_{k+1} dB_t\right) \\ = & \left[\alpha_t + \frac{a_\theta}{1-a_\theta} \Sigma_t \Sigma_t^{-1} (\alpha_t - r_t) - r_t \mathbf{1}_{k+1}\right] dt \\ = & \frac{1}{1-a_\theta} (\alpha_t - r_t \mathbf{1}_{k+1}) dt. \end{aligned}$$

The risk premium with respect to $\mathbb{P}$, dEP, is $(\alpha_t - r_t \mathbf{1}_{k+1})dt$. Therefore, we can find θ by

$$a_\theta = \frac{d\widetilde{EP} - dEP}{d\widetilde{EP}}. \tag{4.17}$$

This relation is similar to the one mentioned in Maenhout [57]. This relation shows that a_θ is proportional to the relative risk premium difference under the different measures.

The market price of risk of a risky asset is the ratio of the risk premium divided by its volatility. Under the measure $\mathbb{P}$, the market price of risk of R_t, given by

$$\lambda_t = \Sigma_t^{-1}(\alpha_t - r_t \mathbf{1}_{k+1}) = \frac{1 - a_\theta}{a_\theta} h_t, \tag{4.18}$$

is unaffected by model uncertainty, following the nominal financial market setting. Under the measure $\widetilde{\mathbb{P}}$, the market price of risk is

$$\widetilde{\lambda}_t = \frac{1}{1 - a_\theta} \Sigma_t^{-1}(\alpha_t - r_t \mathbf{1}_{k+1}) = \frac{1}{a_\theta} h_t. \tag{4.19}$$

This quantity is decreasing in κ. For $\kappa > 0$, $\lambda_t > \widetilde{\lambda}_t$, because $\widetilde{\lambda}_t$ is evaluated under the worst case measure $\widetilde{\mathbb{P}}$. Moreover, a similar relation to (4.17) is also found for λ_t and $\widetilde{\lambda}_t$, i.e., $a_\theta = \frac{\widetilde{\lambda} - \lambda}{\widetilde{\lambda}}$. Last but not least, the second equality in (4.18) and (4.19) provides a way to recover h_t when a_θ and the market price of risk are available.

4.4 The Equilibrium under Model Uncertainty

The previous section incorporates model uncertainty into the investment-consumption decision for a representative agent in a complete market with production, risk-free and risky assets. The utility is a logarithm function of the terminal wealth. In this section, we further investigate the equilibrium implications.

When solving the general optimization problem, the risk-free interest rate r_t, the expected production and risky asset returns in α_t are taken as exogenous. In an equilibrium economy, however, they are determined endogenously by the market clearing condition. Specifically in a representative economic agent settings, this means that the net supply of risky assets and risk-free lending or borrowing must be zero. In other words, for every period in equilibrium, all wealth is invested only in the production improvement, and no wealth is assigned to assets.

Definition 1. *An equilibrium is defined as a set of stochastic processes* $(r_t, \alpha_t, \pi_t, W_t)$ *by*

1. *the market clearing condition,*

$$\pi_t = [\pi_{1t}, \pi'_{2t}]' = [1, \mathbf{0}'_k]' ;$$

2. *the equilibrium terminal wealth* W_T^* *is optimal according to the preference-ordering representation,*

$$\inf_{m_T \in \mathcal{P}_\kappa} \mathbb{E}^{\mathbb{P}} \left[m_T \exp(-bT) \log(W_T) \right] .$$

By assumption, the parameters related to the production process, α_{1t} and σ_{1t} are exogenous. We obtain in Appendix 4.C the closed-form solutions for the market price of risk, the risk-free rate and the risk premium in equilibrium, and summarize them in the following proposition. We use the superscript e to signal the equilibrium-specific parameters and outcomes.

Proposition 3. *Planned at time* τ*, in the presence of model uncertainty, the equilibrium risk-free interest rate is endogenously determined by*

$$r^e_{t,\tau} = \alpha_{1t} - \sigma_{1t}\sigma'_{1t} + a_{\theta^*_\tau}\sigma_{1t}\sigma'_{1t}, \tag{4.20}$$

the expected risky rates by

$$\alpha^e_{2t,\tau} = \alpha_{1t}\mathbf{1}_k - \left(1 - a_{\theta^*_\tau}\right)\left(\sigma_{1t}\sigma'_{1t}\mathbf{1}_k - \sigma_{2t}\sigma'_{1t}\right), \tag{4.21}$$

the risk premium by

$$EP^e_{t,\tau} = \alpha^e_t - r^e_t \mathbf{1}_{k+1} = \left(1 - a_{\theta^*_\tau}\right)\Sigma_t\sigma'_{1t}, \tag{4.22}$$

and the market price of risk by

$$\lambda^e_{t,\tau} = \sigma'_{1t} - a_{\theta^*_\tau}\sigma'_{1t}, \tag{4.23}$$

where $a_{\theta^*_\tau} = -\sqrt{\frac{2\kappa_\tau}{\int_\tau^T \sigma_{1t}\sigma'_{1t}dt}}$*, for a specific model uncertainty amount* $\kappa_{t,\tau}$*.*

As discussed for Proposition 2, the parameter $a_{\theta^*_\tau}$ is planning-time consistent when allowing κ_τ planning-time-varying. As a consequence, the solutions in 3 are implementation-time consistent. The fixed a_{θ^*} is non-positive, and represents the representative agent's conservativeness. The lower it is, the more conservative he is. This attitude is mainly affected by the agent's confidence in the nominal model, or the initial quantity of model uncertainty, κ. When $\kappa \to 0$, $a^*_\theta \to 0$, he is more confident in the nominal model; in the limit, he fully believes in the nominal model. When a_{θ^*} decreases monotonically as κ increases; he has more doubt on the nominal model, and becomes more cautious. Besides, a_{θ^*} depends on the investment horizon T. When T is larger, a_{θ^*} is larger as well, implying a less conservative attitude towards investment. The variance of the production improvement, $\sigma_{1t}\sigma'_{1t}$ has a similar effect as T. This means that when the investment horizon is longer and the production improvement is stable, he would be less and less cautious. At last, a_{θ^*} does not depend on the initial wealth W_0, the expected production return a_1, the risk-free rate r_t, or the discount rate b.

Generally, the equilibrium solutions in Proposition 3 are composed of two parts. One is the standard result without model uncertainty, while the other part stems from model uncertainty. For instance, the standard equilibrium market price of risk is σ'_{1t}, but model uncertainty enlarges the total market price of risk λ^e_t by the amount $-a_{\theta^*}\sigma'_{1t}$. The expected risky rates include an extra part $a_\theta\left(\sigma_{1t}\sigma'_{1t}\mathbf{1}_k - \sigma_{2t}\sigma'_{1t}\right)$. The risk premium contains an extra premium due to model uncertainty given by $-a_\theta\Sigma_t\sigma'_{1t}$. The risk-free rate is affected by an additional amount $a_{\theta^*}\sigma_{1t}\sigma'_{1t}$ as well.

In the presence of model uncertainty, a_{θ^*} is negative. As a result, whether the extra terms will increase or decrease depend on the sign of the value by which a_{θ^*} multiplies. We summarize the changes as follows. Firstly, the risk-free rate decreases in κ, because of the positive variance $\sigma_{1t}\sigma'_{1t}$. Secondly, the change in an expected risky rate depends on whether the variance of the technology return is stronger than the covariance between the technology and the risky asset. If they are negatively correlated, the expected risky rate will definitely decrease in κ. However, if they are positively correlated, the expected risky rate decreases only if the correlation is smaller than the variance of the technology return. Thirdly, the risk premium of the technological return will definitely increase, but the change for each individual asset again depends on the covariance, because $\Sigma_t \sigma'_{1t}$ is in fact a vector containing the correlations of the produced goods with the risky assets.[10] When the production improvement is positively correlated to a risky asset, the risk premium of the risky asset will be increasing, and vice versa. Last but not least, the market price of risk increases in the amount of model uncertainty.

The optimal terminal wealth under the measure $\mathbb{P}$ in equilibrium is given by

$$W_T^{e*} = W_\tau \exp\left\{ \int_\tau^T \left[\alpha_{1t} - \frac{1}{2}\sigma_{1t}\sigma'_{1t} \right] dt + \int_\tau^T \sigma_{1t} dZ_t \right\}.$$

Moreover, the expected terminal wealth is

$$\mathbb{E}^{\mathbb{P}}\left(W_T^{e*}\right) = W_\tau \exp\left(\int_\tau^T \alpha_{1t} dt \right). \tag{4.24}$$

We see that the wealth is not affected by model uncertainty. This is because the wealth is expressed in terms of the physical good which is the numéraire. In equilibrium, the model uncertainty changes both the supply side and the demand side. The investment decision $\pi_t^e = \left[1, \mathbf{0}'_k\right]'$ remains the same for equilibrium . As a result, the model uncertainty does not affect the production as a numéraire, although it affects the prices of assets. If, instead, using the risk-free asset as the numéraire, we would observe the impact of model uncertainty. [11]

[10]It is the first column of $\Sigma_t \Sigma'_t$ that collects the production return's variance $\sigma_{1t}\sigma'_{1t}$ and its correlations with the risky rates $\sigma_{2t}\sigma'_{1t}$.

[11]The rates of the risk-free asset in equilibrium is $B_t = B_\tau \exp\left(\int_\tau^t r_s^e ds\right)$. The expected terminal wealth under the measure $\mathbb{P}$ is

$$\frac{\mathbb{E}^{\mathbb{P}}\left(W_T^{e*}\right)}{B_T} = \frac{W_\tau \exp\left(\int_\tau^T \alpha_{1t} dt\right)}{B_\tau \exp\left(-\int_\tau^t r_s^e ds\right)} = \frac{W_\tau}{B_\tau} \exp\left[\int_\tau^T \left(\sigma_{1t}\sigma'_{1t} - a_{\theta^*}\sigma_{1t}\sigma'_{1t}\right) dt\right].$$

The expected wealth increases in κ, because the risk-free rate r_T^e decreases. The decreasing risk-free rate implies more preference to the present time T over future. When evaluating the wealth with

In Appendix 4.C, the relation between h_t^e and a_θ is found to be

$$h_t^e = a_\theta \sigma'_{1t} = -\sqrt{\frac{2\kappa}{\int_0^T \sigma_{1t}\sigma'_{1t} dt}} \sigma'_{1t}.$$

In the absence of model uncertainty, we have that $\theta = 0$, $a_\theta = 0$, and $h_t^e = 0$, as discussed earlier. Like a_θ, this uncertainty process is related to the volatility of technology improvement; the technology innovations affect the impact of model uncertainty as well.

Given the closed-form solution of h_t^e, it is also interesting to investigate the equilibrium solutions under the alternative measure $\widetilde{\mathbb{P}}$. When evaluated under measure $\widetilde{\mathbb{P}}$, the processes of the production technology Q_t and the risky assets S_t are affected because of the change of dZ by Girsanov's Theorem. But the risk-free asset in equilibrium remains the same as r_t^e. By Girsanov's Theorem, the expected technological return, and the expected risky rates under $\widetilde{\mathbb{P}}$ are, respectively,

$$\begin{aligned}
\widetilde{\alpha}_{1t} =& \alpha_{1t} + \sigma_{1t} h_t^e = \alpha_{1t} + a_{\theta^*} \sigma_{1t} \sigma'_{1t}, \\
\widetilde{\alpha}_{2t}^e =& \alpha_{2t}^e + \sigma_{2t} h_t^e = \alpha_{2t}^e + a_{\theta^*} \sigma_{2t} \sigma'_{1t} \\
=& \alpha_{1t} \mathbf{1}_k - (\sigma_{1t} \sigma'_{1t} \mathbf{1}_k - \sigma_{2t} \sigma'_{1t}) + a_{\theta^*} \sigma_{1t} \sigma'_{1t} \mathbf{1}_k;
\end{aligned}$$

the risk premium is

$$\widetilde{EP}^e = \begin{bmatrix} \widetilde{\alpha}_{1t} \\ \widetilde{\alpha}_{2t}^e \end{bmatrix} - r_t^e \mathbf{1}_{k+1} = \Sigma_t \sigma'_{1t};$$

and the market price of risk is

$$\widetilde{\lambda}_t^e = \sigma'_{1t},$$

satisfying the relation $\widetilde{\lambda}_t^e = \frac{1}{a_\theta} h_t^e$ in (4.19). With some manipulations, one can find that EP^e and $\widetilde{EP}^e$ also satisfy the relation in (4.17). In the general optimization, λ_t under $\mathbb{P}$ is not affected by model uncertainty, while $\widetilde{\lambda}_t$ under $\widetilde{\mathbb{P}}$ decreases in κ. Different in the equilibrium case, λ_t^e under $\mathbb{P}$ increases in κ, while $\widetilde{\lambda}_t^e$ under $\widetilde{\mathbb{P}}$ is

---

the decreasing numéraire of r_T^e, the wealth at time T would be increasing.

unaffected by κ. The main reason is that, in equilibrium, r^e_t and α^e_{2t} are no longer exogenous. Their values, determined endogenously, are affected by κ, and thus so is λ^e_t. $\widetilde{\lambda}^e_t$ is affected because the a^*_θ-related items in $\widetilde{\alpha}_{1t}$ and $\widetilde{\alpha}_{2t}$ are canceled in excess of the risk-free rate r^e_t.

The same quantity of the terminal wealth is expressed by

$$\begin{aligned} W^{e*}_T =& W_0 \exp\left\{ \int_0^T \left[\widetilde{\alpha}_{1t} + \frac{1}{2}\sigma_{1t}\sigma'_{1t} \right] dt + \int_0^T \sigma_{1t} d\widetilde{Z}_t \right\} \\ & W_0 \exp\left\{ \int_0^T \left[\alpha_{1t} + \left(a_{\theta^*} - \frac{1}{2} \right) \sigma_{1t}\sigma'_{1t} \right] dt + \int_0^T \sigma_{1t} d\widetilde{Z}_t \right\}. \end{aligned}$$

This results in a different expected terminal wealth, as compared to the one under the measure $\mathbb{P}$,

$$\mathbb{E}^{\widetilde{\mathbb{P}}}\left(W^{e*}_T\right) = W_0 \exp\left[\int_0^T \left(\alpha_{1t} + a_{\theta^*}\sigma_{1t}\sigma'_{1t} \right) dt \right]. \qquad (4.25)$$

In equilibrium under the measure $\widetilde{\mathbb{P}}$, the expected terminal wealth shows a clear decreasing relation in the model uncertainty due to a^*_θ.

To sum up, in the equilibrium under the measure $\mathbb{P}$, model uncertainty affects the prices of assets but not the numéraire W^{e*}_T. However under the measure $\widetilde{\mathbb{P}}$, the expected return on the exogenous production process $\widetilde{\alpha}_{1t}$ is affected by model uncertainty, and thus so is $\mathbb{E}^{\widetilde{\mathbb{P}}}\left(W^{e*}_T\right)$. Moreover, the expected risky rates $\widetilde{\alpha}_{2t}$ are also affected by model uncertainty by a different amount, as compared to the one in $\mathbb{P}$. However, the risk premium $\widetilde{EP}^e$ and the market price of risk $\widetilde{\lambda}^e_t$ are not affected by model uncertainty.

4.5 Calibrations of the Impact of Model Uncertainty

In this section, we calibrate the parameters, according to the closed-form solutions in equilibrium, to study the impact of model uncertainty on the variables such as expected interest rate, expected risky rates, risk premium, etc. We firstly implement simple calibrations and verify the changes of those closed-form solutions in the amount of model uncertainty κ. Next, we link κ to certain statistical significance levels, and see how much κ can affect those solutions.

4.5.1 Simple Calibrations

In the calibrations, the parameters are assumed to be constant. The production technology process is regarded to represent economic growth, using the monthly data of the OECD year-over-year economic growth rate. We look into two types of risky assets. One is the market portfolio, and the other 5 industry portfolios of consumer goods, manufacture sectors, high-tech sectors, health sectors, and others. Both are retrieved from the Kenneth French data.

We analyze the recent ten years. The reasons of choosing this period are because it is sufficiently long for observation, and the information is up-to date. Moreover, this period includes the financial crisis in 2008, which is one of the sources of model uncertainty. Including it provides us with a natural information to understand the impact of model uncertainty in calibration.

Figure 4.5.1 plots the raw data of the economic growth rate, market portfolio return, and industry portfolios returns from January 2007 to December 2016. As a consequence of the 2008 financial crisis, the market portfolio and the industry portfolios returns dropped to the bottom around October 2008. The DGP growth rate dropped as well, following the financial crisis. The figure shows a half-year lagged effect on the economic growth rate. Considering the lagged effect in the economy, we apply the data of the risky returns from July 2006 to June 2016, and the lagged data of the economic growth rate from January 2007 to December 2016.

Moreover, the economic growth rate data are annualized, while the risky assets' data are expressed in terms of month. For calibration convenience, we adjust the economic growth rate by dividing by 12. Hence, all rates in the following calibrations are expressed in terms of month.

Other values of the parameters for the equilibrium analysis are as follows.

- $\alpha_1 = 0.11\%$. It is the mean of the OECD economic growth rates from year 2007 to 2016, on a monthly basis.

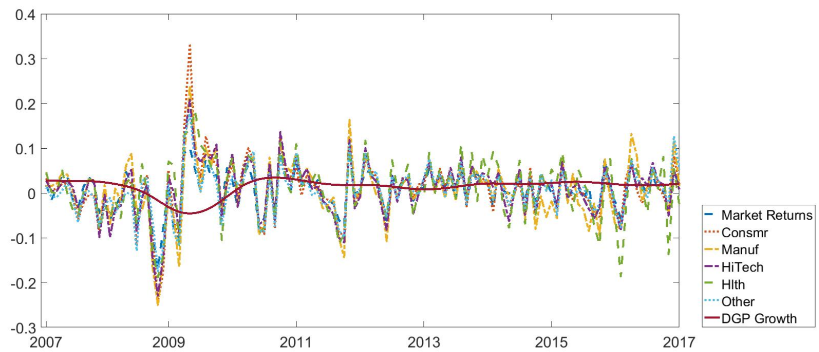


Figure 4.5.1: Economic growth, market portfolio returns and industry portfolios returns.

- σ_1 and σ_2 are the volatilities for the production and risky asset processes. The variance-covariance matrix $\Sigma\Sigma'$ is estimated, providing the variance of the economic growth $\sigma_1\sigma_1'$ and the covariance with the risky assets $\sigma_2\sigma_1'$.

- $b = 0.08333\%$ is the discount rate in terms of month, equivalent to an annualized rate of 1%. It is close to the risk-free rate.

- $W_0 = 1$ is the normalized initial wealth.

- T is the time horizon in months. We calibrate for $T \in [12, 60, 120, 180]$, namely 1 year, 5 years, 10 years, and 15 years.

- k is the number of risky assets. In particular, $k = 1$ or 5.

- κ represents a specific amount of model uncertainty. We investigate the variables of interest against κ. We set the range of κ to be investigated to be $[0, 3.5]$.

Market Portfolio as the Risky Asset

The first calibration uses the market portfolio as the risky asset, i.e., $k = 1$. The return of the market portfolio is used in the analysis. The variance of the production technology return, and the covariance with the expected market return is given by the vector

$$\widehat{\Sigma}\widehat{\sigma}_1' = \begin{bmatrix} 0.0002\% & 0.0022\% \end{bmatrix}'.$$

The variance of the technological improvement $\widehat{\sigma}_1\widehat{\sigma}_1'$, and the covariance $\widehat{\sigma}_2\widehat{\sigma}_1'$ are positive, while the difference between the variance and covariance $\widehat{\sigma}_1\widehat{\sigma}_1' - \widehat{\sigma}_2\widehat{\sigma}_1'$ is negative. According to Proposition 3, this implies that the equilibrium risk-free rate will be decreasing, while the risk premium and the expected market return will be increasing.

Figure 4.5.2 shows the changes against the size of model uncertainty. The value of a_{θ^*} and the equilibrium risk-free rate r^e decrease as κ goes up. In the absence of model uncertainty $\kappa = 0$, $r^e = 0.1116\%$. The risk premium of the market portfolio is increasing in κ, as predicted. When $\kappa = 0$, the expected market return $a_2 = 0.1138\%$, $EP^e = 0.0022\%$. The thin dark dashed lines mark the rates 0.33% and 0.5%, which are simply the annualized equity risk premium of 4% and 6% expressed in terms of months. This shows the possibility to explain the equity risk premium puzzle when sufficient model uncertainty is considered. The expected market return is increasing in κ as well, resulting from the negative difference between the variance of the lagged economic growth and the covariance $(0.0002\% - 0.0022\%)$.

The figures also compare the results when investing over different investment time horizons. In general, the variables are more sensitive to the amount of model uncertainty when investing in shorter time horizons. This implies that for the same level of uncertainty, the impact on the market is stronger when the decision is made over a short period; the representative agent tends to behave more conservatively and cautiously when his investment only lasts for a short time, and thus the market becomes more conservative as well, driving up the expected risky rates and the risk premium and pressing down the risk-free rates. In our calibrations, for $T = 12$ month investment horizon, the equilibrium interest rate falls below 0 when more than $\kappa = 3.14$ amount of model uncertainty is taken into account; the threshold of κ for the zero expected interest rate is larger for the longer investment horizon.

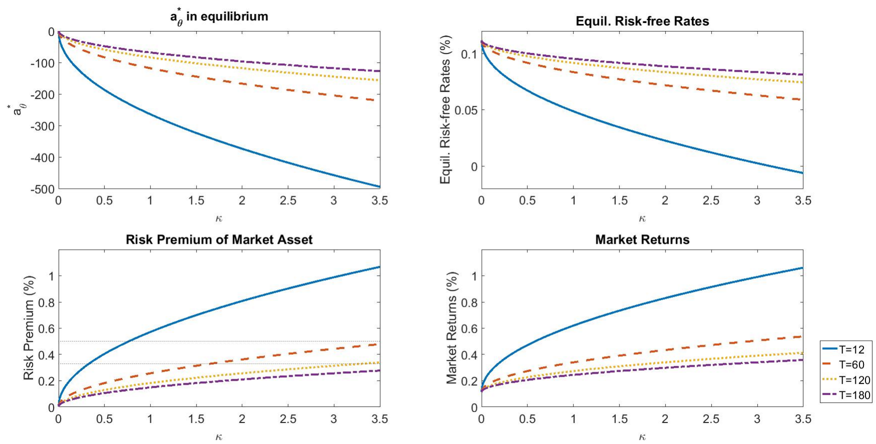


Figure 4.5.2: The equilibrium results using market portfolio as the risky asset. The parameter a_{θ^*} and equilibrium risk-free rate decrease in κ. Both the equity risk premium and the expected market return increase. At $\kappa = 3.14$, the equilibrium risk-free rate for $T = 12$ becomes negative.

5 Industry Portfolios as the Risky Assets

In the second calibration, we use the $k = 5$ industry portfolios as the risky assets, whose covariances with the lagged economic growth are collected in the vector

$$\widehat{\Sigma}\widehat{\sigma}_1' = \begin{bmatrix} 0.0002\% & 0.0016\% & 0.0023\% & 0.0021\% & 0.0012\% & 0.0028\% \end{bmatrix}'.$$

By Proposition 3, this vector shows that the risk premium will increase. Moreover, the most sensitive one to model uncertainty is the fifth portfolio, because of the largest covariance 0.0028% can cause the strongest strength to change in (4.22). The forth portfolio is the least sensitive one among the 5 portfolios. The portfolio expected returns will be increasing, because of the negative difference between $\sigma_1\sigma_1'$ and $\sigma_1\sigma_2'$. The calibration results are shown in Figure 4.5.3.

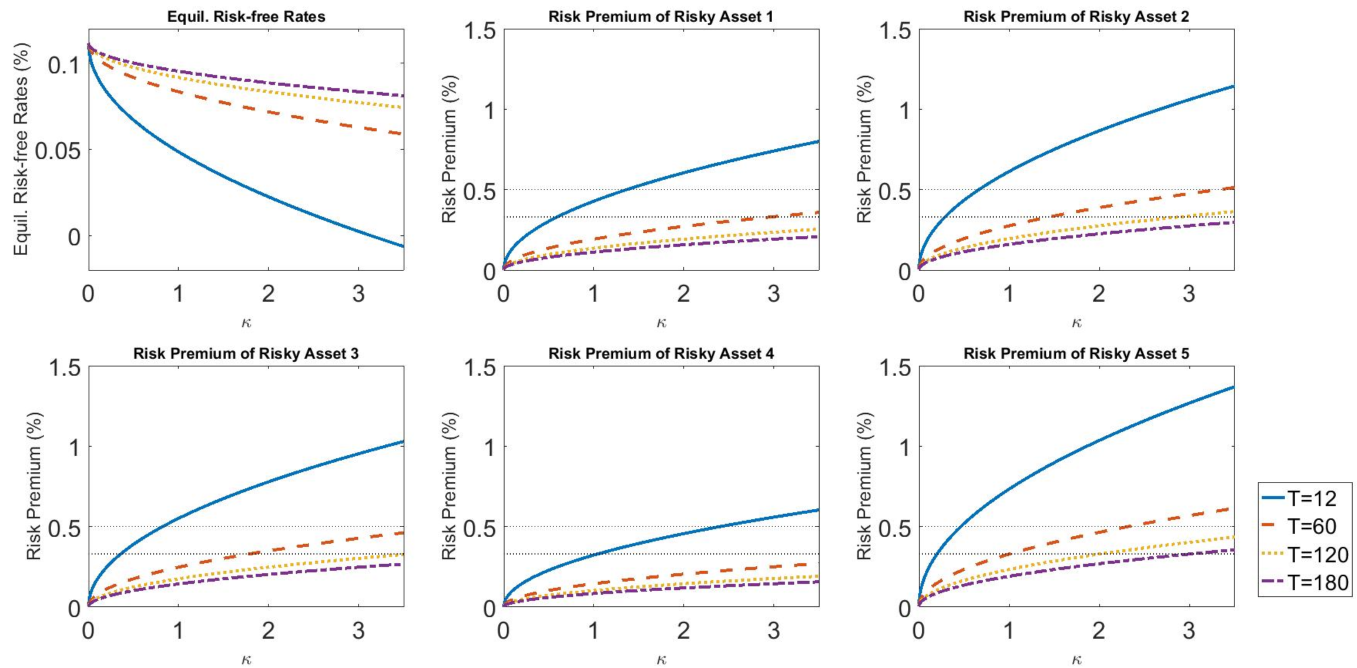


Figure 4.5.3: The equilibrium risk-free rate and the risk premium of the 5 industry portfolios.

The Expected Terminal Wealth and Expected Utility

We look into the expected terminal wealth and the expected utility in equilibrium under the measure $\widetilde{\mathbb{P}}$, which are affected by model uncertainty.

In the following investigation, we use the market portfolio as the risky asset. Figure 4.5.4 plots the changes. Resulting from the model uncertainty, the expected terminal wealth is reduced under the alternative measure $\widetilde{\mathbb{P}}$, and accordingly so is the expected utility.

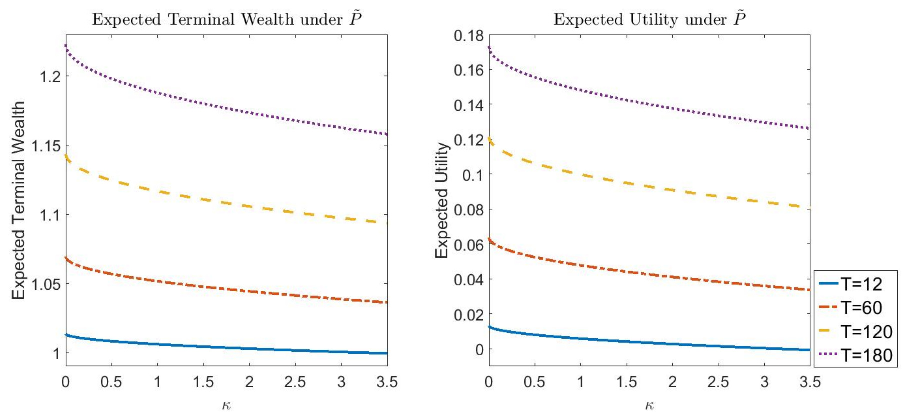


Figure 4.5.4: The optimal terminal wealth in equilibrium and the corresponding expected utility under measure $\widetilde{\mathbb{P}}$, given different investment horizon T. The risky asset is the market portfolio.

4.5.2 A Reasonable κ

The calibrations above illustrate the changes in the significant variables against the model uncertainty amount κ, as shown in Proposition 3. In general, the larger κ is, the less confident the representative agent is in the nominal model, and behaves more conservatively. This causes the changes. A closely related question is then raised: how large should a reasonable κ be? This subsection will address this question.

The amount of model uncertainty is implied by the size of the uncertainty set. A common approach applied in the literature to determine the size of the uncertainty set is the detection error probability approach [2]; see applications in Gagliardini et al. [24], Maenhout [57, 58]. It is an approach based on the empirical data and the likelihood ratio, and aims to form a constraint on the uncertainty set to contain only statistically indistinguishable models. In other words, this approach takes into account parameter uncertainty only. Maenhout [57] calibrates a reasonable κ by this approach and finds that the equity risk premium can be explained by his calibrated κ.

Alternatively, one may consider an approach pertaining to the uncertainty set construction of (4.6). The approach is based on a test statistic such that the model is statistically indistinguishable if the KL divergence is less than the test value of certain confidence level of $1-\alpha$. According to the Theorem 9.1 in Pardo [61], the uncertainty set of $1-\alpha$ confidence level is constructed by

$$m_T \in \mathcal{P}_\kappa = \left\{ m_T \middle| \mathbb{E}\left(m_T \log m_T\right) \le \kappa, \quad \kappa = \frac{\chi^2_{dof,1-\alpha}}{2n} \right\},$$

where $\chi^2_{dof,1-\alpha}$ is the χ^2 distribution with degree of freedom dof and the critical value α. The parameter n is the number of observations. The degree of freedom is the number of unknown variables dof, i.e, the number of parameters used to parametrized $\alpha_t(\theta)$ and $\Sigma_t(\theta)$. The value of dofspecifically given by $(k+1)+(k+1)^2$. A higher n or lower k would result in a lower κ, meaning that more observations or less risky assets will shrink the size of the model uncertainty set containing only statistically indistinguishable models.

In response to the calibrations, we look into the κ for the number of asset $k=1$ and $k=5$, respectively. Table 4.5.1 reports the reasonable κ that contains indistinguishable models, for the sample size $n=[30, 50, 100, 120]$ and the critical value $\alpha=[10\%, 5\%, 1\%]$, respectively.

In Table 4.5.1, κ reduces for large sample size and less risky assets, given the $1-\alpha$ significance level. This is understandable because the parameter uncertainty decreases in the sample size and in the dimension. Moreover, for a larger significance level, κ is larger yielding a bigger model uncertainty set containing more statistically indistinguishable models.

Table 4.5.1: A reasonable κ with $k = 1$ and $k = 5$ risky assets

n	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 1\%$
		Panel A: $k = 1$ risky assets	
30	0.1774	0.2099	0.2802
50	0.1064	0.1259	0.1681
100	0.0532	0.0630	0.0841
120	**0.0444**	**0.0525**	**0.0700**
		Panel B: $k = 5$ risky assets	
30	0.9015	0.9687	1.1034
50	0.5409	0.5812	0.6621
100	0.2705	0.2906	0.3310
120	**0.2254**	**0.2422**	**0.2759**

In the Equilibrium

We compare the κ values in the table and Figure 4.5.2 for market portfolio calibration. Although, incorporating model uncertainty indeed helps in explaining the risk premium puzzle for $T = 12$, it is still far from explaining the equity risk premium puzzle if more periods are taken into account. A similar situation happens for some of the five risky portfolios.

In the calibrations, we use 10 years monthly data, which includes 120 observations. The corresponding κ considering only parameter uncertainty are marked boldfaced. It can be easily observed that given the κ in Table 4.5.1, it is almost impossible to explain the equity risk premium puzzle. In order to resolve the puzzle, we have to allow for a much larger κ. In other words, model misspecification uncertainty is necessary to explain the 4% to 6% annualized equity risk premium.

Moreover, the correlation between the economic growth and risky asset affects as well the size of κ in explaining the equity risk premium puzzle. When the correlation is low, it implies more model misspecification uncertainty is needed. The same implication applies to longer investment horizon.

In the General Solution

For the general optimization problem, the closed-form solution for θ^* in (4.13) is negative only if the integration part in the denominator is larger than κ. We apply the data to calculate the value of the integration part for $T = 1$. We find that it is 268 and 2052 for $k = 1$ and $k = 5$, respectively. When $T > 1$, these values will be even larger. Compare the values the reasonable values of κ in Table 4.5.1. We believe that the numerator in the solution (4.13) is positive in reality, and thus $\theta^* < 0$.

4.6 Summary and Conclusions

This paper incorporates model uncertainty into a modified Cox et al. [11] model. The model considers an investment-consumption choice problem in a complete market setting. The expected utility is a logarithm function of the terminal consumption, which exhausts all wealth without bequest at the last period. The dynamics of the state variables in the original Cox et al. [11] settings is not considered in this paper.

We propose an alternative approach to Hansen and Sargent [34] to obtain time consistent solution. Our approach is attractive as it avoid the complexity of structuring a discount RN derivative in Hansen and Sargent [34]. We are able to find, in a much easier way, a fixed and time independent discount uncertainty parameter that indicates the strength of the impact of model uncertainty.

We firstly find the portfolio rules for the investment-consumption decision. The impact of model uncertainty is embedded in the investment decision, increasing the demand of the risk-free asset and decreasing the terminal wealth. The results also imply an effectively increased risk aversion. Moreover, we establish the link between the discount uncertainty parameter and the uncertainty process. Based on this relation, we explore the connections between different measures for the market price of risk and the equity risk premium, respectively. Our finding is similar to that in Maenhout [57]. Also, we find that both the market price of risk and the equity risk premium decrease in the amount of model uncertainty when evaluated under the alternative measure. They are unaffected when evaluated under the nominal measure.

We further find the asset pricing rules through the equilibrium, where all wealth is reinvested in improving the production technology. We derive the closed-form solutions for the expected risk-free rate, the expected risky rates, the expected equity risk premium and the market prices of risk. The solutions suggest decreasing risky-free rate and increasing market prices of risk against the amount of model uncertainty. The changes of the expected risky returns and the risk premiums depend on the correlation between the production and the risky assets. The expected wealth measured under the nominal measure is unaffected by model uncertainty. Also important to point out is that a longer investment horizon and less volatility of the production technology improvement could relax the conservativeness towards investment, and affect the market as if less model uncertainty is taken into account.

At last, we calibrate to analyze the impact of model uncertainty on asset pricing. The calibrations show that the expected equilibrium risk-free rates could decrease below 0. This helps in explaining the risk-free rate puzzle and the negative interest rates. Moreover, the calibrations show that the risk premium increases because of the positive correlation between the production improvement and risky assets, providing a possibility to explain the risk premium puzzle. However, for some assets with small correlation, only considering parameter uncertainty is not enough to resolve the puzzle. We show that it is necessary to take in account the model misspecification uncertainty, in order to fully explain the puzzle. This is not completely in line with the finding in Maenhout [57, 58], who applies Anderson et al. [2]'s framework that only allows for parameter uncertainty and finds it sufficient to explain the puzzle in his calibrations.

Appendix

Appendix 4.A Assumptions

Following Glasserman and Xu [27], in order to solve the optimization problem (4.8) by swapping the optimizers $\sup_{\pi_t}$ and $\sup_\theta$, we make the following assumptions.

1. Π, the parameter set for π_t is compact. This is true for π_t because the investment weight is in the interval $[0, 1]$.

2. The original utility function $\exp\left[-b(T-\tau)\right]\log\left(W_T\right)$ is concave. Expanding the function gives

$$
\begin{aligned}
&\exp\left[-b(T-\tau)\right]\log\left(W_T\right) = \\
&\exp\left[-b(T-\tau)\right]\Bigg[\log W_\tau + \int_\tau^T \pi'_{t,\tau}\Sigma_t dZ_t \\
&+ \int_\tau^T \left(\pi'_{t,\tau}\left(\alpha_t - r_t\mathbf{1}_{k+1}\right) + r_t - \frac{1}{2}\pi'_{t,\tau}\Sigma_t\Sigma'_t\pi_{t,\tau}\right) dt\Bigg],
\end{aligned}
$$

a function of $\pi_{t,\tau}$. If $\pi_{t,\tau} \in \mathbb{R}$, we have

$$
\begin{aligned}
\frac{\partial}{\partial \pi_{t,\tau}} \exp\left[-b(T-\tau)\right]\log\left(W_T\right) &> 0, \\
\text{and} \frac{\partial^2}{\partial \pi_{t,\tau}^2} \exp\left[-b(T-\tau)\right]\log\left(W_T\right) &< 0,
\end{aligned}
$$

implying the concavity of the function in $\pi_{t,\tau}$.

3. The moment generating function $F(\theta) = \mathbb{E}\left[\exp\left(\theta \exp\left[-b(T-\tau)\right]\log\left(W_T\right)\right)\right]$ exists for θ in some open set containing the origin. This ensures the finiteness of $\hat{F}(\theta)$ so that the optimal solution of m_T

$$
m_T^* = \frac{\exp\left(\theta \exp\left[-b(T-\tau)\right]\log\left(W_T\right)\right)}{\mathbb{E}\left[\exp\left(\theta \exp\left[-b(T-\tau)\right]\log\left(W_T\right)\right)\right]}
$$

is well-defined.

4. If $\left(\theta^*, \pi^*_{t,\tau}, m^*_T\right)$ solves the problem, $\theta^*_\tau \in [\theta^*_{\min,\tau}, 0]$ for some $\theta^*_{\min,\tau} \in [-\infty, 0]$ such that for any $\theta_\tau \in [\theta^*_{\min,\tau}, 0]$, the set $\left\{\pi_{t,\tau} \in \Pi : \hat{F}(\theta_\tau) > -\infty\right\}$ is compact.

5. For any $\theta_\tau \in [\theta^*_{\min,\tau}, 0]$, $\mathbb{E}\left[\exp\left(\theta_\tau\left[-b(T-\tau)\right]\log\left(W_T\right)\right)\right]$ is lower semi-continuous in $\pi_{t,\tau}$.

Appendix 4.B Solutions for the General Case

This appendix solves the inner optimization problem of (4.10) subject to (4.5) by the marginal approach, and then the outer optimization with respect to the uncertainty parameter θ_τ.

Applying the martingale approach, the problem (4.10) is simplified to be an optimization of a power utility of terminal wealth

$$\sup_{\pi_{t,\tau}} -\mathbb{E}\left(W_T^{a_{\theta_\tau}}\right),$$

subject to

$$\mathbb{E}\left[H_{T,\tau} W_T\right] = W_\tau \tag{4.26}$$

where $H_{t,\tau}$ is the state density process, or the kernel pricing process, defined by

$$H_{t,\tau} := \frac{L_t}{B_t}.$$

L_t is in fact an RN-derivative to change the risky measure to the risk neutral measure, satisfying the stochastic differential equation

$$dL_t = -L_t \zeta_t' dZ_t, \qquad L_\tau = 1, \tag{4.27}$$

where, when the market is complete, by the risky asset process and Girsanov Theorem $\zeta_t \in \mathbb{R}^{(k+1)\times 1}$ has an explicit form

$$\zeta_t = \Sigma_t^{-1}(\alpha_t - r_t \mathbf{1}_{k+1}) \tag{4.28}$$

and known as the market price of risk process. B_t is the money market price, following the risk-free asset process

$$dB_t = r_t B_t dt.$$

Consider the Lagrangian function

$$\mathcal{L} := -W_T^{a_{\theta_\tau}} + \lambda \left(H_{T,\tau} W_T - W_\tau \right).$$

Take the first order condition with respect to W_T,

$$-a_{\theta_\tau} W_T^{a_{\theta_\tau}-1} = -\lambda H_{T,\tau},$$

and it yields the optimal

$$W_T^* = \left(\frac{\lambda H_{T,\tau}}{a_{\theta_\tau}} \right)^{\frac{1}{a_{\theta_\tau}-1}}. \tag{4.29}$$

Substituting the optimal W_T^* in the budget constraint (4.26)

$$\left(\frac{\lambda}{a_{\theta_\tau}} \right)^{\frac{1}{a_{\theta_\tau}-1}} \mathbb{E}\left[\left(H_{T,\tau} \right)^{\frac{a_{\theta_\tau}}{a_{\theta_\tau}-1}} \right] = W_\tau,$$

from which we find

$$\lambda = \frac{a_{\theta_\tau} W_\tau^{a_{\theta_\tau}-1}}{\left\{ \mathbb{E}\left[\left(H_{T,\tau} \right)^{\frac{a_{\theta_\tau}}{a_{\theta_\tau}-1}} \right] \right\}^{a_{\theta_\tau}-1}}. \tag{4.30}$$

Substituting λ of (4.30) back to W_T^* of (4.29) gives the explicit solution

$$W_T^* = \frac{W_\tau H_{T,\tau}^{\frac{1}{a_{\theta_\tau}-1}}}{\mathbb{E}\left[H_{T,\tau}^{\frac{a_{\theta_\tau}}{a_{\theta_\tau}-1}}\right]}$$

Also, $H_{t,\tau}$ can be obtained, given by $H_{t,\tau} = H_{\tau,\tau} \exp\left[\int \left(-r_t - \frac{1}{2}\zeta_t'\zeta_t\right) dt - \int \zeta_t' dZ_t\right]$. By stochastic Leibnitz Rule, it be proved that the budget constraint (4.31) is a martingale process because it follows

$$d\left(H_{t,\tau}W_t\right) = H_{t,\tau}W_t(\pi_{t,\tau}'\Sigma_t - \zeta_t')dZ_t. \tag{4.31}$$

Therefore, we can write

$$H_{t,\tau}W_t = \mathbb{E}_t\left[H_{T,\tau}W_T\right] = W_\tau \frac{\mathbb{E}_t\left[H_{T,\tau}^{\frac{a_{\theta_\tau}}{a_{\theta_\tau}-1}}\right]}{\mathbb{E}\left[H_{T,\tau}^{\frac{a_{\theta_\tau}}{a_{\theta_\tau}-1}}\right]}. \tag{4.32}$$

One may write $H_{t,\tau}^{\frac{a_{\theta_\tau}}{a_{\theta_\tau}-1}} = g_t\Lambda_t$ where

$$\begin{aligned} g_t &:= \exp\left\{\frac{a_{\theta_\tau}}{1-a_{\theta_\tau}}\int_\tau^t r_u du + \frac{a_{\theta_\tau}}{2(1-a_{\theta_\tau})^2}\int_\tau^t \zeta_u'\zeta_u du\right\} \\ \Lambda_t &:= \exp\left\{\frac{a_{\theta_\tau}}{1-a_{\theta_\tau}}\int_\tau^t \zeta_u' dZ_u - \frac{(a_{\theta_\tau})^2}{2(1-a_{\theta_\tau})^2}\int_\tau^t \zeta_u'\zeta_u du\right\} \end{aligned}$$

If r_u and ζ_u are deterministic, then g_t is deterministic and Λ_t is a martingale by construction. So

$$\mathbb{E}_t\left[H_{u,\tau}^{\frac{a_{\theta_\tau}}{1-a_{\theta_\tau}}}\right] = g_u\Lambda_t, \quad \tau \le t \le u \le T.$$

Define $N_t := g_t$. We have $\mathbb{E}\left[H_{T,\tau}^{\frac{a_{\theta_\tau}}{a_{\theta_\tau}-1}}\right] = N_T$ because $\Lambda_\tau = 1$. Then (4.32) becomes

$$G_t = H_{t,\tau}W_t = \frac{W_\tau g_T \Lambda_t}{N_T}.$$

Moreover,

$$dG_t = \frac{W_\tau m_T}{N_T} d\Lambda_t = \frac{W_\tau m_T \Lambda_t}{N_T} \frac{a_{\theta_\tau}}{1 - a_{\theta_\tau}} \zeta_t' dZ_t = H_\tau(t) W_t \frac{a_{\theta_\tau}}{1 - a_{\theta_\tau}} \zeta_t' dZ_t \quad (4.33)$$

Comparing (4.31) and (4.33), we have

$$\pi_{t,\tau}' \Sigma_t - \zeta_t' = \frac{a_{\theta_\tau}}{1 - a_{\theta_\tau}} \zeta_t'.$$

Consequently, the optimal solution of investment is given by

$$\pi_{t,\tau}^* = \frac{1}{1 - a_{\theta_\tau}} \Sigma_t'^{-1} \zeta_t = \frac{1}{1 - a_{\theta_\tau}} (\Sigma_t \Sigma_t')^{-1} (\alpha_t - r_t \mathbf{1}_{k+1}),$$

and the optimal terminal wealth is given by

$$\begin{aligned} W_T^* &= W_\tau \exp\left\{ \int_\tau^T \left[\pi_{t,\tau}^{*\prime} (\alpha_t - r_t \mathbf{1}_{k+1}) + r_t - \frac{1}{2} \pi_{t,\tau}^{*\prime} \Sigma_t \Sigma_t' \pi_{t,\tau}^* \right] dt + \int_\tau^T \pi_{t,\tau}^{*\prime} \Sigma_t dZ_t \right\} \\ &= W_\tau \exp\left\{ \int_\tau^T \left[\frac{1 - 2a_{\theta_\tau}}{2(1 - a_{\theta_\tau})^2} (\alpha_t - r_t \mathbf{1}_{k+1})' (\Sigma\Sigma')^{-1} (\alpha_t - r_t \mathbf{1}_{k+1}) + r_t \right] dt \right. \\ &\qquad \left. + \int_\tau^T \frac{1}{1 - a_{\theta_\tau}} (\alpha_t - r_t \mathbf{1}_{k+1})' \Sigma'^{-1} dZ_t \right\}, \end{aligned}$$

By substituting in the optimization problem (4.10), we have

$$\begin{aligned} &\sup_{\theta_\tau < 0} J(\theta_\tau) + \frac{\kappa_\tau}{\theta_\tau} = \sup_{\theta_\tau < 0} \frac{1}{\theta_\tau} \log \mathbb{E}_\tau \left[\exp\left(a_{\theta_\tau} \log W_T^* \right) \right] + \frac{\kappa_\tau}{\theta_\tau} \\ &= \sup_{\theta_\tau < 0} X_{T,\tau} \log W_\tau + \int_\tau^T X_{T,\tau} \left[\frac{1}{(1 - \theta_\tau X_{T,\tau})} (\alpha_t - r_t \mathbf{1}_{k+1})' (\Sigma_t \Sigma_t')^{-1} (\alpha_t - r_t \mathbf{1}_{k+1}) + r \right] dt + \frac{\kappa_\tau}{\theta_\tau}, \end{aligned}$$

given $a_{\theta_\tau} = \theta_\tau X_{T,\tau}$. The optimization problem is equivalent to

$$\sup_{\theta_\tau < 0} \int_\tau^T \frac{1}{(1 - \theta_\tau X_{T,\tau})} (\alpha_t - r_t \mathbf{1}_{k+1})' (\Sigma_t \Sigma_t')^{-1} (\alpha_t - r_t \mathbf{1}_{k+1}) X_{T,\tau} dt + \frac{\kappa_\tau}{\theta_\tau}.$$

Taking the first order condition

$$(1-\theta_\tau X_{T,\tau})^{-2}\int_\tau^T (\alpha_t - r_t\mathbf{1}_{k+1})'(\Sigma_t\Sigma_t')^{-1}(\alpha_t - r_t\mathbf{1}_{k+1})X_{T,\tau}^2 dt - \frac{\kappa_\tau}{\theta_\tau^2} = 0,$$

or

$$\left[\int_\tau^T (\alpha_t - r_t\mathbf{1}_{k+1})'(\Sigma\Sigma')^{-1}(\alpha_t - r_t\mathbf{1}_{k+1})dt - \kappa_\tau\right] X_{T,\tau}^2\theta_\tau^2 + 2\kappa_\tau X_{T,\tau}\theta_\tau - \kappa_\tau = 0$$

gives the optimal θ_τ^*

$$\theta_\tau^* = \frac{-\kappa_\tau - \sqrt{\kappa_\tau \int_\tau^T (\alpha_t - r_t\mathbf{1}_{k+1})'(\Sigma\Sigma')^{-1}(\alpha_t - r_t\mathbf{1}_{k+1})dt}}{\left[\int_\tau^T (\alpha_t - r_t\mathbf{1}_{k+1})'(\Sigma\Sigma')^{-1}(\alpha_t - r_t\mathbf{1}_{k+1})dt - \kappa_\tau\right] X_{T,\tau}}.$$

Appendix 4.C Solutions for the Equilibrium Case

Combining the general solution of the optimal portfolio choice in (4.11), we establish the following relation,

$$\begin{bmatrix} 1 \\ \mathbf{0}_k \end{bmatrix} = \pi_t^{e*} = \frac{1}{1-a_\theta}\left(\Sigma_t\Sigma_t'\right)^{-1}\left(\alpha_t^e - r_t^e\mathbf{1}_{k+1}\right). \tag{4.34}$$

From (4.34), the equilibrium interest rate related to the product process is given by

$$r_t^e = \alpha_{1t} - (1-a_\theta)\,\sigma_{1t}\sigma_{1t}',$$

and related to risky assets is given by

$$r_t^e\mathbf{1}_k = \alpha_{2t} - (1-a_\theta)\,\sigma_{2t}\sigma_{1t}'.$$

From this two equations, we find the returns on the risky assets in equilibrium,

$$\alpha_{2t}^e = \alpha_{1t}\mathbf{1}_k - \sigma_{1t}\left[(1-a_\theta)\,\sigma_{1t}'\right]\mathbf{1}_k + \sigma_{2t}\left[(1-a_\theta)\,\sigma_{1t}'\right]$$

The market price of risk in equilibrium can also be derived, given by

$$\lambda_t^e = \begin{bmatrix} \sigma_{1t} \\ \sigma_{2t} \end{bmatrix}^{-1} \begin{bmatrix} \alpha_{1t} - r_t^e \\ \alpha_{2t}^e - r_t^e \mathbf{1}_k \end{bmatrix} = (1 - a_\theta)\, \sigma_{1t}'.$$

As a whole, the equilibrium interest rate can be simply expressed as

$$r_t^e = \alpha_{1t} - \sigma_{1t}\lambda_t^e$$

Recall the wealth dynamic process under the measure $\mathbb{P}$,

$$W_T^* = W_\tau \exp\left\{ \int_\tau^T \left[\pi_t^{*'} (\alpha_t - r_t \mathbf{1}_{k+1}) + r_t - \frac{1}{2} \pi_t^{*'} \Sigma_t \Sigma_t' \pi_t^* \right] dt + \int_\tau^T \pi_t^{*'} \Sigma_t dZ_t \right\}$$

In equilibrium, $\pi^{e*} = [1, \mathbf{0}_k']'$. Thus,

$$W_T^{e*} = W_\tau \exp\left[\int_\tau^T \left(\alpha_{1t} - \frac{1}{2} \sigma_{1t}\sigma_{1t}' \right) dt + \int_\tau^T \sigma_{1t} dZ_t \right],$$

is unaffected by model uncertainty under the measure $\mathbb{P}$. So is the expected terminal wealth in equilibrium.

Remembering $a_{\theta,\tau} = \theta_\tau X_{T,\tau}$, the optimization problem (4.10) for equilibrium case is then

$$\begin{aligned} &\sup_{\theta_\tau < 0} \frac{1}{\theta_\tau} \log \mathbb{E}^{\mathbb{P}} \left(W_T^{*e \theta_\tau X_{T,\tau}} \right) + \frac{\kappa_\tau}{\theta_\tau} \\ = &\sup_{\theta_\tau < 0} X_T \log W_0 + X_T \int_\tau^T \left[\alpha_{1t} - \frac{1}{2} (1 - \theta_\tau X_{T,\tau}) \sigma_{1t}\sigma_{1t}' \right] dt + \frac{\kappa_\tau}{\theta_\tau}, \end{aligned}$$

wherein

$$\begin{aligned} \mathbb{E}^{\mathbb{P}} \left(W_T^{e* \theta X_{T,\tau}} \right) =& \mathbb{E}^{\mathbb{P}} \left\{ W_\tau^{\theta_\tau X_{T,\tau}} \exp\left[\theta_\tau X_{T,\tau} \int_\tau^T \left(\alpha_{1t} - \frac{1}{2} \sigma_{1t}\sigma_{1t}' \right) dt + \theta X_{T,\tau} \int_\tau^T \sigma_1 dZ_t \right] \right\} \\ =& W_\tau^{\theta_\tau X_{T,\tau}} \exp\left[\theta_\tau X_{T,\tau} \int_\tau^T \left(\alpha_{1t} - \frac{1}{2} \sigma_{1t}\sigma_{1t}' \right) dt + \frac{1}{2} \theta_\tau^2 X_{T,\tau}^2 \int_\tau^T \sigma_{1t}\sigma_{1t}' dt \right]. \end{aligned}$$

equivalent to

$$\sup_{\theta_\tau>0} \frac{1}{2}\theta X^2_{T,\tau} \int_\tau^T \sigma_{1t}\sigma'_{1t}dt + \frac{\kappa_\tau}{\theta_\tau}.$$

It shows that the initial wealth and production return do not play a role in the optimization problem. Also, the optimization problem itself is a static problem, easy to be solved by the first order condition (FOC) given by

$$f(\theta_\tau) = \frac{1}{2}X^2_{T,\tau} \int_\tau^T \sigma_{1t}\sigma'_{1t}dt - \frac{\kappa_\tau}{\theta^2_\tau} = 0.$$

Recalling $X_{T,\tau} = \exp\left[-b(T-\tau)\right]$, the optimal solution for θ_τ is given by

$$\theta^*_\tau = -\sqrt{\frac{2\kappa_\tau}{\exp\left[-2b(T-\tau)\right]\int_\tau^T \sigma_{1t}\sigma'_{1t}dt}},$$

for $\theta^*_\tau < 0$, and

$$a_{\theta^*_\tau} = -\sqrt{\frac{2\kappa_\tau}{\exp\left[-2b(T-\tau)\right]\int_\tau^T \sigma_{1t}\sigma'_{1t}dt}}\exp\left[-b(T-\tau)\right] = -\sqrt{\frac{2\kappa_\tau}{\int_\tau^T \sigma_{1t}\sigma'_{1t}dt}}.$$

As discussed, $a_{\theta^*_\tau}$ is planning time consistent for a properly determined κ_τ. And since $a_{\theta^*_\tau}$ is planning time consistent, it can be determined by the initial amount of model uncertainty κ.

Next, we look for the h_t in terms of a_θ. According to the relation between (4.9) and (4.14), we can establish the relation for the equilibrium that gives

$$\int_0^T -\frac{1}{2}a^2_\theta\sigma_{1t}\sigma'_{1t}dt + \int_0^T a_\theta\sigma_{1t}dZ_t = \int_0^T h^{e'}_t dZ_t - \int_0^T \frac{1}{2}h^{e'}_t h^e_t dt, \tag{4.35}$$

The diffusion parts of (4.35) imply that

$$h^e_t = a_\theta\sigma'_{1t} + A_t, \tag{4.36}$$

where $A'_t Z_t = 0$. If and only if $A_t = 0$ could (4.36) be in line with the implication from the drift terms of (4.35), i.e.,

$$h_t^{e'} h_t^e = a_\theta^2 \sigma_{1t} \sigma_{1t}'.$$

The wealth process under the measure $\widetilde{\mathbb{P}}$ is transformed from

$$W_T^{e*} = W_0 \exp\left\{ \int_0^T \left[\pi_t^{*'} (\alpha_t + \Sigma_t h_t - r_t \mathbf{1}_{k+1}) + r_t - \frac{1}{2} \pi_t^{*'} \Sigma_t \Sigma_t' \pi_t^* \right] dt + \int_0^T \pi_t^{*'} \Sigma_t dZ_t \right\}.$$

Given $h_t^e = a_\theta \sigma_{1t}'$, thus the terminal wealth given by

$$W_T^{e*} = W_0 \exp\left\{ \int_0^T \left[\alpha_{1t} + \left(a_\theta - \frac{1}{2} \right) \sigma_{1t} \sigma_{1t}' \right] dt + \int_0^T \sigma_{1t} d\widetilde{Z}_t \right\},$$

and the corresponding expected utility given by

$$\mathbb{E}^{\widetilde{\mathbb{P}}} \left[\exp(-bT) \log \left(W_T^{e*} \right) \right] = \exp(-bT) \left[\log W_0 + \int_0^T \left(\alpha_{1t} + a_\theta \sigma_{1t} \sigma_{1t}' \right) dt \right],$$

are affected by model uncertainty.

References

[1] T. Adrian, R. K. Crump, and E. Moench. Pricing the term structure with linear regressions. Journal of Financial Economics, 110(1):110 – 138, 2013.

[2] E. W. Anderson, L. P. Hansen, and T. J. Sargent. A quartet of semigroups for model specification, robustness, prices of risk, and model detection. Journal of the European Economic Association, 1(1):68–123, 03 2003.

[3] R. Bansal and A. Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. The Journal of Finance, 59(4):1481–1509, 2004.

[4] A. Ben-Tal, D. den Hertog, A. D. Waegenaere, B. Melenberg, and G. Rennen. Robust solutions of optimization problems affected by uncertain probabilities. Management Science, 59(2):341–357, 2013.

[5] T. Breuer and I. Csiszár. Measuring distribution model risk. Mathematical Finance, 2013.

[6] D. Brigo and F. Mercurio. Interest rate models- Theory and Practice. Springer Finance. Springer, Berlin, 2nd ed edition, 2006.

[7] M. Cagetti, L. P. Hansen, T. Sargent, and N. Williams. Robustness and pricing with uncertain growth. Review of Financial Studies, 15(2):363–404, 2002.

[8] Z. Chen and L. Epstein. Ambiguity, risk, and asset returns in continuous time. Econometrica, 70(4):1403–1443, July 2002.

[9] J. H. Cochrane and M. Piazzesi. Bond risk premia. The American Economic Review, 95(1):pp. 138–160, 2005.

[10] J. H. Cochrane and M. Piazzesi. Decomposing the yield curve. Working paper, March 2008.

[11] J. C. Cox, J. E. Ingersoll, and S. A. Ross. An intertemporal general equilibrium model of asset prices. Econometrica, 53(2):pp. 363–384, 1985.

[12] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. Econometrica, 53(2):pp. 385–407, 1985.

[13] Q. Dai and K. Singleton. Term structure dynamics in theory and reality. Review of Financial Studies, 16(3):631–678, 2003.

[14] Q. Dai and K. J. Singleton. Specification analysis of affine term structure models. The Journal of Finance, 55(5):1943–1978, 2000.

[15] F. X. Diebold and C. Li. Forecasting the term structure of government bond yields. Journal of Econometrics, 130(2):337 – 364, 2006.

[16] F. X. Diebold, G. D. Rudebusch, and S. B. Aruoba. The macroeconomy and the yield curve: a dynamic latent factor approach. Journal of Econometrics, 131(1–2):309 – 338, 2006.

[17] J. Driessen, B. Melenberg, and T. Nijman. Common factors in international bond returns. Working paper, Tilburg University, 2000.

[18] G. R. Duffee. Information in (and not in) the term structure. The Review of Financial Studies, 24(9):2895–2934, 2011.

[19] D. Duffie and R. Kan. A yield-factor model of interest rates. Mathematical Finance, 6(4):379–406, 1996.

[20] D. Ellsberg. Risk, ambiguity, and the savage axioms. The Quarterly Journal of Economics, 75(4):643–669, 1961.

[21] L. G. Epstein and M. Schneider. Recursive multiple-priors. Journal of Economic Theory, 113(1):1 – 31, 2003.

[22] C. Friedman. Confronting model misspecification in finance: Tractable collections of scenario probability measures for robust financial optimization problems. International Journal of Theoretical and Applied Finance, 05(01): 33–54, 2002.

[23] P. Gagliardini, P. Porchia, and F. Trojani. Ambiguity aversion, bond pricing and the non-robustness of some affine term structures. Working paper, February 2005.

[24] P. Gagliardini, P. Porchia, and F. Trojani. Ambiguity aversion and the term structure of interest rates. Review of Financial Studies, 22(10):4157–4188, 2009.

[25] L. Garlappi, R. Uppal, and T. Wang. Portfolio selection with parameter and model uncertainty: A multi-prior approach. Review of Financial Studies, 20 (1):41–81, 2007.

[26] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. Journal of Mathematical Economics, 18(2):141 – 153, 1989.

[27] P. Glasserman and X. Xu. Robust risk measurement and model risk. Quantitative Finance, 14(1):29–58, 2014.

[28] S. Goncalves and H. White. Bootstrap standard error estimates for linear regression. Journal of the American Statistical Association, 100(471):970–979, 2005.

[29] M. Guidolin and F. Rinaldi. Ambiguity in asset pricing and portfolio choice: a review of the literature. Theory and Decision, 74(2):183–217, 2013.

[30] R. S. Gürkaynak, B. Sack, and J. H. Wright. The U.S. treasury yield curve: 1961 to the present. Journal of Monetary Economics, 54(8):2291 – 2304, 2007.

[31] J. D. Hamilton and J. C. Wu. Identification and estimation of gaussian affine term structure models. Journal of Econometrics, 168(2):315 – 331, 2012.

[32] L. P. Hansen and T. J. Sargent. Acknowledging misspecification in macroeconomic theory. Review of Economic Dynamics, 4(3):519 – 535, 2001.

[33] L. P. Hansen and T. J. Sargent. Robust Control and Model Uncertainty. American Economic Review, 91(2):60–66, May 2001.

[34] L. P. Hansen and T. J. Sargent. Robust estimation and control under commitment. Journal of Economic Theory, 124(2):258 – 301, 2005.

[35] L. P. Hansen and T. J. Sargent. Recursive robust estimation and control without commitment. Journal of Economic Theory, 136(1):1 – 27, 2007.

[36] L. P. Hansen and T. J. Sargent. Robustness. Princeton University Press, 2008.

[37] L. P. Hansen and T. J. Sargent. Fragile beliefs and the price of uncertainty. Quantitative Economics, 1(1):129–162, 2010.

[38] L. P. Hansen and T. J. Sargent. Robustness and ambiguity in continuous time. Journal of Economic Theory, 146(3):1195 – 1223, 2011.

[39] L. P. Hansen, T. J. Sargent, and J. Tallarini, Thomas D. Robust permanent income and pricing. The Review of Economic Studies, 66(4):pp. 873–907, 1999.

[40] L. P. Hansen, T. J. Sargent, and N. E. Wang. Robust permanent income and pricing with filtering. Macroeconomic Dynamics, 6:40–84, 2 2002. ISSN 1469-8056.

[41] L. P. Hansen, T. J. Sargent, G. Turmuhambetova, and N. Williams. Robust control and model misspecification. Journal of Economic Theory, 128(1): 45 – 90, 2006.

[42] P. R. Hansen, A. Lunde, and J. M. Nason. The model confidence set. Econometrica, 79(2):453–497, 2011.

[43] N. Hari, A. De Waegenaere, B. Melenberg, and T. E. Nijman. Longevity risk in portfolios of pension annuities. Insurance: Mathematics and Economics, 42(2):505–519, 2008.

[44] G. Iyengar and A. K. C. Ma. A robust optimization approach to pension fund management. Journal of Asset Management, 11, 2010.

[45] S. Joslin, K. J. Singleton, and H. Zhu. A new perspective on gaussian dynamic term structure models. Review of Financial Studies, 2011.

[46] S. Joslin, M. Priebsch, and K. Singleton. Risk premiums in dynamic term structure models with unspanned macro risks. The Journal of Finance, 69 (3):1197–1233, 2014.

[47] N. Ju and J. Miao. Ambiguity, learning, and asset returns. Econometrica, 80 (2):559–591, 2012.

[48] I. Karatzas and S. E. Shreve. Methods of mathematical finance. Applications of mathematics. Springer, New York, 1998.

[49] F. Knight. Risk, Uncertainty and Profit. Houghton Miffin, Boston, 1921.

[50] S. Kullback and R. A. Leibler. On information and sufficiency. The Annals of Mathematical Statistics, 22(1):pp. 79–86, 1951.

[51] J. Li. An empirical investigation of model uncertainty, with an application to affine term structure models. Doctoral dissertation, Tilburg University, 2017.

[52] R. B. Litterman and J. Scheinkman. Common factors affecting bond returns. The Journal of Fixed Income, 1(1):54–61, 1991.

[53] H. Liu. Robust consumption and portfolio choice for time varying investment opportunities. Annals of Finance, 6(4):435–454, 2010.

[54] S. C. Ludvigson and S. Ng. Macro factors in bond risk premia. Review of Financial Studies, 22(12):5027–5067, 2009.

[55] F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica, 74(6):1447–1498, 2006.

[56] F. Maccheroni, M. Marinacci, and A. Rustichini. Dynamic variational preferences. Journal of Economic Theory, 128(1):4 – 44, 2006.

[57] P. J. Maenhout. Robust portfolio rules and asset pricing. Review of Financial Studies, 17(4):951–983, 2004.

[58] P. J. Maenhout. Robust portfolio rules and detection-error probabilities for a mean-reverting risk premium. Journal of Economic Theory, 128(1):136 – 163, 2006.

[59] R. C. Merton. An Intertemporal Capital Asset Pricing Model. Econometrica, 41(5):867–87, September 1973.

[60] C. R. Nelson and A. F. Siegel. Parsimonious modeling of yield curves. The Journal of Business, 60(4):473–489, 1987.

[61] L. Pardo. Statistical Inference Based on Divergence Measures. CRC Press, Abingdon, 2005.

[62] A. Pelsser. Pricing in incomplete markets. Working paper, 2011.

[63] M. Piazzesi. Bond yields and the federal reserve. Journal of Political Economy, 113(2):pp. 311–344, 2005.

[64] M. Piazzesi. Affine term structure models. In Handbook of Financial Econometrics, chapter 12. Elsevier, 2010.

[65] J. C. Schneider and N. Schweizer. Robust measurement of (heavy-tailed) risks: Theory and implementation. Journal of Economic Dynamics and Control, 61:183 – 203, 2015.

[66] S. Shen, A. Pelsser, and P. C. Schotman. Robust hedging in incomplete markets. 2014.

[67] L. E. Svensson. Estimating and interpreting forward interest rates: Sweden 1992 - 1994. Working Paper 4871, National Bureau of Economic Research, September 1994.

[68] F. Trojani and P. Vanini. Robustness and ambiguity aversion in general equilibrium. Review of Finance, 8(2):279, 2004.

[69] O. Vasicek. An equilibrium characterization of the term structure. Journal of Financial Economics, 5(2):177 – 188, 1977.

[70] A. Vissing-Jørgensen and O. P. Attanasio. Stock-market participation, intertemporal substitution, and risk-aversion. The American Economic Review, 93(2):383–391, 2003.

[71] J. H. Wright. Term premia and inflation uncertainty: Empirical evidence from an international panel dataset. The American Economic Review, 101 (4):1514–1534, 2011.